The Dissertation Committee for Andrei Khilkevich

Certifies that this is the approved version of the following dissertation:

CEREBELLAR MECHANISMS UNDERLYING ADAPTIVE MOTOR RESPONSES

Committee:

Michael D. Mauk, Supervisor

Alexander Huk

Jonathan Pillow

Hiroshi Nishiyama

Richard Aldrich

CEREBELLAR MECHANISMS UNDERLYING ADAPTIVE MOTOR RESPONSES

by

Andrei Khilkevich

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2017

Acknowledgements

This dissertation would not be possible without the support from many people. Going chronologically, I should start with thanking all the talented and passionate Physics teachers I was lucky to encounter early on. Particularly, to Ilja Zabelinsky and Mikhail Seniuk. Without witnessing your drive and genuine curiosity towards science, I would have not likely ended up doing research and being where I am today.

I want to give special thanks to my advisor, Mike Mauk. I enrolled at UT without any background in neuroscience, and in fact did not even join the Neuroscience program at first. Thank you Mike for introducing me to the field of neuroscience and for making this transition easier than it could have been. Your intelligence, critical thinking and amazing intuition have shaped the way I think scientifically. Thank you for the freedom to try out all my crazy ideas. Thanks to all of my committee members: Alex Huk, Hiroshi Nishiyama, Rick Aldrich and Jonathan Pillow - I really appreciate all the comments and feedback you provided over the years.

Thanks to all the talented undergraduates that I was lucky to work with. Loc, Juan, Molly-Marie, Nicole, Ashkan, Biraj – the behavioral data presented here would be impossible to collect without your help. To Hunter and Frank – thank you for introducing me, a theoretical physicist at that time, to the realm of experiments in neuroscience. Hunter – thank you for being there to bounce around ideas for experiments as well as for teaching me about in-vivo recordings and tetrode drives. I think that after

mastering how to work with tetrodes I can manually assemble pretty much anything. Neto, Sami and Evan – thank you for being in the lab, you made the experience more worthwhile.

To the members of Ila Fiete's lab, John and Rishi, the opportunity to discuss analysis ideas and to learn from you two was invaluable. Also, John, thank you for taking that biophysics class. If I did not meet you there, chances are I would still be doing particle Physics. Special thanks to Boris Zemelman for patiently tolerating the requests of viral constructs (in rabbit-scale volumes) for all my optogenetic experiment attempts. Lastly – thanks to other past and present lab members: Loren, Chris, Joe, Yujin, Wen, Rose, Vyash – I learned a great deal from you all.

CEREBELLAR MECHANISMS UNDERLYING ADAPTIVE MOTOR RESPONSES

Andrei Khilkevich, Ph.D

The University of Texas at Austin, 2017

Supervisor: Michael D. Mauk

Abstract

In order to understand the function of any brain structure, one must know what input/output transformation it performs. The term input/output transformation includes at least two stages. First, we must understand how inputs are processed. Second, we must know what the output activity encodes. Certain properties of the cerebellum make such an undertaking feasible. In this thesis I present the results of three main projects designed to study the input/output transformations of this major brain system from different angles.

In the first project I investigated the relationship between spiking activity of cerebellar cortex principal neurons - Purkinje cells (PCs) - and eyelid conditioned response (CRs) profiles on a single trial basis. Systematically exploring a variety of encoding possibilities, I found that PCs do not directly encode a single kinematic variable

of a CR. The best prediction was rather achieved via a dynamical model approach, where PCs provide a 'drive' to the eyelid plant, the dynamics of which are described by a differential equation.

In the second project I addressed how the cerebellum deals with inherent uncertainty about the nature of sensory inputs. I found that under conditions of uncertainty, the cerebellum performed a probabilistic binary choice, scaling the probability of response with the similarity between current and trained stimuli. Importantly, if responses were made, their amplitude was close to the previously trained value, maintaining the adaptive nature of responses. Recordings from eyelid Purkinje cells localized this computation to cerebellar cortex. Results from large-scale computer simulation suggest that the efference copy signal is critical for the expression of target response amplitude.

In the third project I studied cerebellar mechanisms of learning and expression of movement sequences. While the majority of movements we perform are composed of sequences, most of the knowledge about cerebellar learning and computation comes from tasks involving single, unitary movements. Hence, I designed a novel sequence training protocol to explicitly test the ability of the cerebellum to chain together a series of movements through associative learning processes. The results demonstrate a simple yet general framework for how the cerebellum can learn to produce a movement sequence.

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 General overview

The vast knowledge of cerebellar anatomy and physiology has greatly facilitated the ability to design and interpret results of experiments studying the cerebellum. The majority of cerebellar cell types and synaptic organization have been outlined about half a century ago (Eccles, 1967; Ito, 1984), although some additions to the cerebellar synaptic connectivity are made almost every year (Hull and Regehr, 2012; Houck and Person, 2014; Kim et al., 2014; Ankri et al., 2015). Fig. 1.1. A illustrates the connectivity between different cells types in the cerebellum. In large, the cerebellum can be divided into two spatially separated regions: cerebellar cortex and deep cerebellar nuclei (DCN). Cerebellar cortex forms an outer layer of folded gray matter while cerebellar nuclei form distinct clusters of neurons residing in the white matter. Anatomically, cerebellar cortex is divided into 10 distinct lobules. Structurally however, the cerebellar cortex appears to have largely the same uniform connectivity throughout different lobules (Ito, 1984). Such uniformity of cerebellar structure and connectivity allows the translation of computational principles found from a specific behavioral paradigm to general cerebellar phenomena.

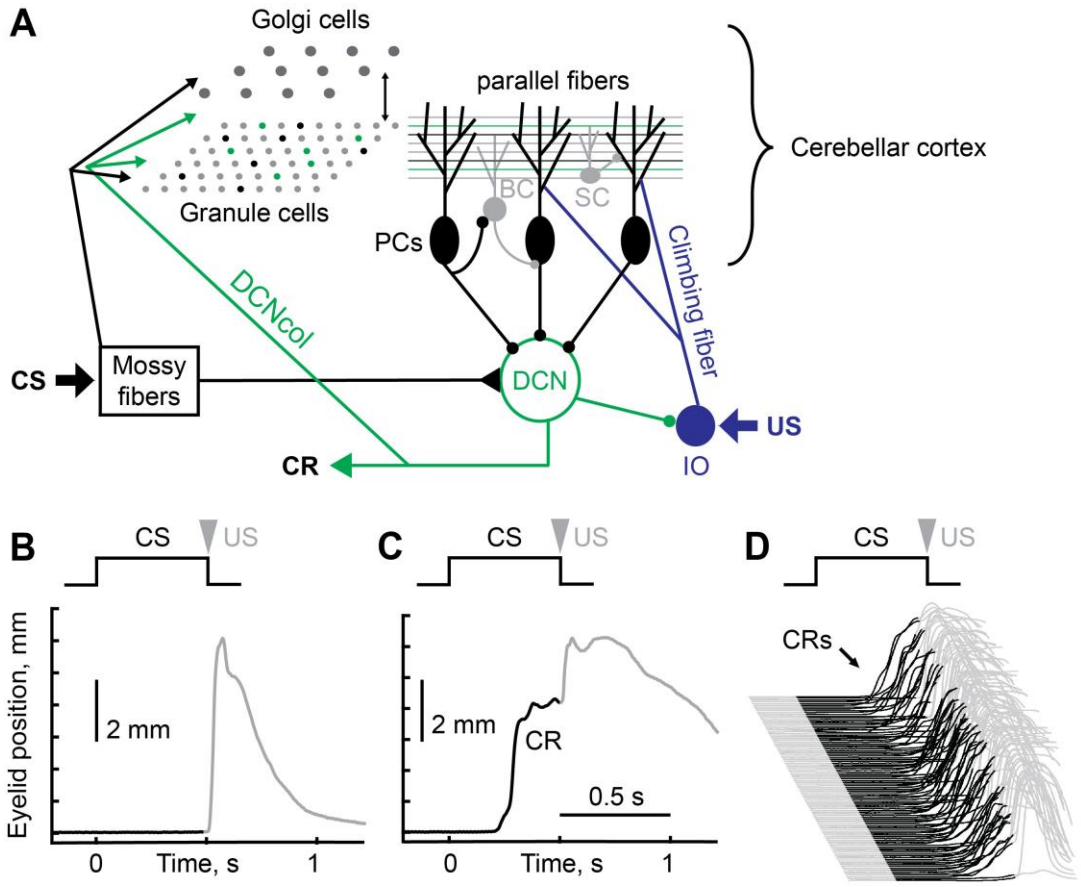

Fig. 1.1. Schematics of cerebellar circuitry and eyelid conditioning paradigm.
A) Schematic diagram of cerebellar circuitry. Inhibitory connections are indicated by circles at the end of the line. B-D Eyelid conditioning paradigm. Stimulus presentation is shown at the top, eyelid position as a function of time is shown below. Upward deviation corresponds to eyelid closure. B) Neutral conditioned stimulus (CS) is paired with eyelid stimulation as unconditioned stimulus (US). In naïve state, there is no eyelid response during CS, US evokes a reflexive non-cerebellar eyelid closure (shown in grey). C) Behavioral response after animal learned CS-US pairing. Predictive eyelid closure during CS and before US is a conditioned response (CR). D) Waterfall plot showing behavioral responses during eyelid conditioning session. Trials are arranged chronologically from bottom to top.

All external inputs to the cerebellum arrive via two afferent routes: mossy fibers and climbing fibers. Mossy fibers originate from pontine nuclei neurons (Ito, 1984; Shinoda et al., 2000) that receive a variety of multimodal sensory information (Boyd and Aitkin, 1976; Maekawa et al., 1981) as well as inputs from several areas of cerebral cortex (Brodal, 1968; Glickstein et al., 1980; Cicirata et al., 2005; Leergaard et al., 2006). Through mossy fibers the cerebellum receives information about current sensory inputs and the state of other brain areas. On their way to the cerebellum mossy fibers pass through the middle cerebellar peduncle and branch into two routes: a portion of mossy fiber collaterals form direct synaptic connections with neurons in DCN and another branch projects to the cerebellar cortex. In the cerebellar cortex mossy fibers synapse onto: 1) granule cells – numerous excitatory interneurons, 2) Golgi cells – inhibitory interneurons and 3) unipolar brush cells – excitatory interneurons. From a computational and connectivity perspective, all three cell types can be considered the input processing stage of the cerebellar cortex. Only the axons of granule cells – parallel fibers – project to other types of neurons in the cerebellar cortex. The most prominent targets of parallel fibers are Purkinje cells (PCs), the sole output neurons of the cerebellar cortex. Parallel fibers form by different estimates from 80000 (Palkovits et al., 1971) to 200000 (Eccles, 1967) synapses per single PC. From early theoretical work (Marr, 1969; Albus, 1971), it has been suggested that plasticity at parallel fibers-to-PCs synapses can mediate learning in the cerebellar cortex. Since then numerous experiments have verified the existence of postsynaptic LTD and LTP at this synapse (Otmakhov et al., 1997; Hansel and Linden,

2000; Gall, 2005), established the rules for plasticity induction (Safo and Regehr, 2005, 2008) and confirmed its necessity for cerebellar motor learning (Hansel et al., 2006). Apart from PCs, parallel fibers synapse onto Golgi, basket and stellate cells. The latter two types of inhibitory interneurons are located in the molecular layer of cerebellar cortex and inhibit PCs. PCs have somewhat unusual properties compared to principal neurons in other brain areas. First, PCs have unusually high background firing rate – around 50 Hz in-vivo. This high firing rate is partially generated by spontaneous activity of PCs themselves and partially by excitatory input provided through parallel fibers by granule cells. Second, PCs are inhibitory neurons and therefore provide a strong tonic inhibition to DCN neurons downstream.

Climbing fibers, which originate from neurons in the inferior olive (Desclin, 1974; Shinoda et al., 2000), convey an input to the cerebellum that controls behavioral learning (McCormick et al., 1985; Mauk et al., 1986; Medina et al., 2002) and synaptic plasticity (Safo and Regehr, 2008; Suvrathan et al., 2016). In the cerebellar cortex climbing fibers form vast synaptic connections along the dendritic trees of PCs. While a single climbing fiber contacts several PCs, a given PC receives input from only a single climbing fiber (Kano and Hashimoto, 2009; Carrillo et al., 2013). Such specificity is achieved during the developmental stage (Hashimoto et al., 2009; Carrillo et al., 2013) and is likely to be crucial for cerebellar learning. Climbing fiber input to PCs controls bidirectional plasticity at parallel fiber to PCs synapses (Hansel and Linden, 2000; Safo and Regehr, 2008). Therefore the climbing fiber is thought to provide "a teaching signal"

(Marr, 1969; Mauk and Ohyama, 2004) to guide plasticity in the cerebellar cortex. Because of the massive innervation of the PC dendritic tree, each climbing fiber spike causes a reliable action potential in the target PC. Due to opening of voltage-gated calcium channels and consequent calcium influx, the action potential waveform is more prolonged and therefore referred to as a "complex spike" (Simpson et al., 1996; Lang et al., 1999; Ohmae and Medina, 2015) as opposed to other normal simple spikes. PC's complex spikes can be recorded both intracellularly (Eilers et al., 1995; Gao et al., 2012) and extracellularly (Jirenhed et al., 2007; Gao et al., 2012; Halverson et al., 2015). This electrophysiological signature provides a useful way to identify PCs during extracellular recordings and allows restricting analysis to PCs with complex spikes evoked by a specific stimulus. I will utilize both of these advantages later in the manuscript. Inferior olive neurons have low, but non-zero, background firing rate around 1 Hz in-vivo (Keating and Thach, 1995). This low spontaneous activity is thought to be necessary to prevent synaptic weights at parallel fibers to PC synapses from drifting with time (Kenyon, 1997; Mauk and Donegan, 1997; Kenyon et al., 1998).

The sole output of the cerebellum is conveyed via DCN neurons. GABAergic DCN neurons are known to inhibit inferior olive neurons (Best and Regehr, 2009; Lefler et al., 2014) while glutamatergic cells project to brainstem and midbrain nuclei, thalamus and spinal cord (Flumerfelt et al., 1973; Asanuma et al., 1983a, 1983b). Through the thalamus the cerebellum projects, among other areas of cerebral cortex, to premotor and primary motor cortexes (Kelly and Strick, 2003; Proville et al., 2014) and is thought to

modulate movement execution via influence on primary motor cortex activity. A notable exception is the interpositus nucleus which is a portion of DCN. The Interpositus nucleus projects to the red nucleus (Flumerfelt et al., 1973) which in turn projects to the facial nucleus. Facial nucleus neurons are lower motor neurons that innervate muscles controlling facial movements. Therefore, a cerebellar output can directly result in motor movements if the movement is facial (though this is not the only possible scenario). Such a direct relationship makes it possible to infer cerebellar output from a simple behavioral readout. A behavioral paradigm which we used in experiments within this manuscript – eyelid conditioning – utilizes this advantage.

1.2 Eyelid conditioning paradigm as a tool for studying the cerebellum

Eyelid conditioning is a form of classical conditioning that contributed greatly to knowledge about learning (Garcia et al., 1999; Medina and Mauk, 1999; Ohyama and Mauk, 2001; Khilkevich et al., 2016), timing (Buonomano and Mauk, 1994; Medina et al., 2000; Kalmbach et al., 2010a), plasticity rules and sites (Medina and Mauk, 1999; Medina et al., 2001; Ohyama et al., 2006; Voicu and Mauk, 2006; Lee et al., 2015) and general computational mechanisms in the cerebellum (Kenyon et al., 1998; Kalmbach et al., 2011). One of the main reasons for such vast and productive usage of eyelid conditioning is that it engages the cerebellum in a very direct way. In eyelid conditioning (Fig. 1.1 B-D), repeated pairings of a neutral conditioned stimulus (CS) followed by an

unconditioned stimulus (US), which evokes a non-cerebellar reflexive closure of the eye, results in subjects learning to produce an anticipatory eyelid closure, which is referred to as a conditioned response (CR). Figure 1.1 B demonstrates a trial during early acquisition where CR I not produced and US evokes reflexive eyelid closure. Following panel C demonstrates a trial later in acquisition with CR present. Panel D shows behavioral responses during an example session. I would like to note that although CRs are produced by eyelid closure, they are not reflexive blinks (Schade Powers et al., 2010). A gradual change in CR amplitude is apparent during acquisition and extinction sessions. Moreover a study has shown that the cerebellum can be trained to produce CRs with a target amplitude (Kreider and Mauk, 2010). A variety of sensory stimuli can be used as a CS: auditory (Mauk and Ruiz, 1992; Halverson et al., 2010), visual (Halverson et al., 2009) or a direct electrical stimulation of lateral geniculate nucleus (Halverson et al., 2009), superior colliculus (Halverson et al., 2009), visual cortex (Halverson et al., 2009) or medial auditory thalamus (Campolattaro et al., 2007; Halverson and Freeman, 2010). In each case the CS is conveyed to the cerebellum through mossy fibers. This provides a powerful approach to study the cerebellum by controlling temporal characteristics of its input via direct electrical stimulation of mossy fibers (Steinmetz et al., 1986; Svensson et al., 2010; Kalmbach et al., 2011). In this way a CS is also fully restricted to the cerebellum and contributions of upstream sensory areas to any observed phenomena can be excluded. For the reasons above, in the majority of experiments presented in the manuscript I used electrical stimulation of mossy fibers as a CS. A US is conveyed to the

cerebellum via climbing fibers and can be replaced by electrical stimulation of inferior olive neurons (Mauk et al., 1986; Steinmetz et al., 1989).

Early lesion studies have identified two regions in the cerebellum that are necessary for the acquisition and expression of CRs. A portion of DCN, the anterior interpositus nucleus (AIN), is necessary for acquisition and drives CR expression (Krupa and Thompson, 1995, 1997). AIN neurons are known to increase activity during CRs (McCormick and Thompson, 1984a; Halverson et al., 2010), and AIN lesions abolish the expression of CRs (McCormick and Thompson, 1984b; Yeo et al., 1985a; Chapman et al., 1990). Lesions of cerebellar cortex severely disrupt timing of CRs and prevent both acquisition and extinction (Yeo et al., 1985b; Garcia and Mauk, 1998; Kalmbach et al., 2010a). PCs, the sole output of cerebellar cortex, show decreases in activity during CRs (Hesslow and Ivarsson, 1994; Jirenhed et al., 2007; Halverson et al., 2015; ten Brinke et al., 2015), and optogenetic silencing of PCs immediately elicits movement (Heiney et al., 2014).

The vast majority of studies also agree that no plasticity relevant to eyelid conditioning is stored downstream from AIN. A support for this notion comes from studies where pharmacological inactivation of either superior cerebellar peduncle (Krupa and Thompson, 1995) or brainstem motor nuclei, including facial nucleus, (Krupa et al., 1996) was performed during acquisition sessions. Such manipulation robustly blocked the expression of CRs (and URs in case of facial nucleus inactivation) during acquisition

sessions. However, once the inactivation was removed, CRs immediately appeared as opposed to the gradual acquisition observed under normal conditions. These results, combined with the fact that inactivation of AIN prevents learning and expression of CRs (Krupa et al., 1993), argue that regions downstream from AIN do not participate in learning eyelid CRs.

Such a direct relationship between the activity of PCs or AIN neurons and the motor response makes it reasonable to seek how cerebellar output neurons encode the full trajectory of eyelid CRs. **In Chapter 2** I seek to find a precise relationship between eyelid PCs activity and behavior on a single trial basis. Several studies have in some way addressed this question, but they have been limited by: 1) assuming a linear relationship between neuronal activity and movement (Medina and Lisberger, 2009; Heiney et al., 2014); 2) using only some movement characteristics (e.g. amplitude), but not the full trajectory; 3) assuming a fixed time delay between firing rate and movement while ignoring the contribution from the spike history (Medina and Lisberger, 2007, 2009). The majority of studies (Medina and Lisberger, 2007, 2009) have also used an inverse model approach, which have provided important insights, but cannot be used to directly predict motor output. I present results of a systematic analysis using a linear-nonlinear model–based approach (Serruya et al., 2002; Paninski et al., 2004; Hochberg et al., 2006) applied on data from a published dataset (Halverson et al., 2015). This approach overcomes the previously mentioned limitations as it incorporates spike-history dependence along with empirically-based non-linearity. I demonstrate that PC activity does not specifically

encode any single kinematic variable of the eyelid CR (though the assumption of linear velocity encoding gives reasonably accurate results, consistent with published studies (Medina and Lisberger, 2007, 2009; Heiney et al., 2014). In contrast, the best prediction accuracy is achieved with a dynamical model. There, the dynamics of a muscle plant (eyelid in our case) is described by the second-order linear differential equation that resembles the Newton's second law, and PC activity provides a drive (external force term in the equation). I demonstrate that eyelid PC activity is sufficient to capture a large portion of the trial-to-trial behavioral variance (both in CR amplitude and CR onset time). In addition I explore the population code of PCs by studding how prediction accuracy of the eyelid CR profile scales with the number of PCs used, demonstrating that activity of a few PCs (even a single PC) is sufficient to achieve high prediction accuracy. In subsequent chapters I therefore use eyelid PCs activity as the main measure of cerebellar cortex output.

In Chapter 3 I report a novel adaptation that the cerebellum uses in the face of noisy and unreliable inputs to keep the output adaptive. When a subject is trained in an eyelid conditioning paradigm, it usually is put into a dim, sound-attenuating box. The idea is to prevent any external stimuli from interfering with the CS while the animal is being trained or expressing CRs. This approach, used traditionally in classical conditioning, results in robust learning and has contributed greatly to our knowledge about different brain systems. It is also, however, intrinsically limited.

In the real world, the cerebellum needs to process incomplete, noisy inputs and their combinations. Even after successful learning to a particular input, a new input that is only partially similar to the learned input creates an ambiguous situation. On the one hand, partial similarity could arise from the original learned input being contaminated by noise. On the other hand, there is a chance that it is a new, novel input to which the cerebellum has not learned any response yet. Using a combination of behavioral experiments, single unit recordings from eyelid PCs and large-scale computer simulations, I demonstrate that the cerebellum possesses a computational adaptation that maintains adaptive motor and network responses even in the face of such ambiguity.

In Chapter 4 I explore how the cerebellum implements learning and execution of movement sequences. The majority of movements that we perform on a daily basis are not singular but are rather comprised of sequences of movements. Studies of sequence learning in human patients with cerebellar lesions revealed either severe deficits (Doyon et al., 1997; Shimansky et al., 1997) or an inability to learn the sequence at all (Shin and Ivry, 2003). These and other studies suggest that the cerebellum is involved in learning of motor sequences. However, no studies have shown explicitly how learning of motor sequences is implemented in the cerebellum. To address this question, I designed a novel eyelid conditioning training protocol to explicitly test the ability of the cerebellum to chain together a series of movements through an associative learning process. Results demonstrate a simple yet general framework for how the cerebellum can learn to generate a sequence of appropriately timed responses.

In sum, the topics covered in this manuscript address crucial aspects of cerebellar computation: (I) transformation of cerebellar output activity into a motor response, (II) processing of ambiguous or noisy inputs and (III) cerebellar implementation of movement sequences.

1.3 General Methods

1.3.1 Surgery

In all experiments subjects were New Zealand albino rabbits (Oryctolagus cuniculus; Myrtle's Rabbitry) weighing 2.5-3.5 kg at experiment onset. Treatment of rabbits and surgical procedures were in accordance with National Institutes of Health guidelines and an institutionally approved animal welfare protocol. All subjects were maintained on a 12 h light/dark cycle. One week before the start of each experiment, subjects were removed from the home cage and anesthetized with a cocktail of acepromazine (1.5 mg/kg) and ketamine (45 mg/kg). After onset of anesthesia, the subjects were placed in a stereotaxic frame, intubated, and maintained on isoflurane (1~2% mixed in oxygen) for the remainder of the surgery. Under sterile conditions the skull was exposed with a midline incision (~5 cm), and four holes were drilled for anchor screws. Some anchor screws also functioned as ground screws for subjects with mossy

fiber stimulation implants or a microdrive ground for subjects with a microdrive implant. The rabbit's head was then positioned with lambda 1.5 mm ventral to bregma.

For subjects prepared only for behavioral experiments involving electrical stimulation of mossy fibers, a craniotomy was drilled out at 5.5 mm lateral and 3 mm anterior from lambda, ipsilateral to the trained eye. Skull fragments were carefully removed from the craniotomies, the dura matter was carefully opened under visual guidance. One or two laterally spaced (by 1 mm) tungsten stimulating electrodes (A-M Systems, Carlsborg, WA; tip exposed to obtain impedance of 100–200 kΩ) were implanted in the middle cerebellar peduncle (16 mm ventral to lambda).

For subjects prepared for microdrive implantation in the cerebellar cortex, a craniotomy was also drilled out at 5.9 mm posterior and 6.0 mm lateral to lambda. Skull fragments were carefully removed from the craniotomies, the dura matter was carefully opened under visual guidance. A custom-made microdrive (16 tetrodes and 2 references) fitted with an electronic interface board (EIB-36-16TT, Neuralynx) was implanted in the left anterior lobe of the cerebellar cortex at a 40° angle posterior to vertical and 17.8 mm ventral to lambda. This region of the anterior lobe has been shown to be involved in acquisition and expression of well-timed conditioned eyelid responses (Garcia et al., 1999; Kalmbach et al., 2010b; Halverson et al., 2015). The primary target of tetrode recordings were PCs with evoked complex spikes from the US, referred to throughout the manuscript as eyelid PCs. The bundle cannula of the microdrive was lowered to the

surface of the brain. The craniotomies were sealed with low viscosity silicon (Kwik-Sil; World Precision Instruments). Ahead bolt to mount the eyelid detector, anchor/ground screws, stimulation electrodes and microdrive were secured with dental acrylic (Bosworth Fastray, Pink; The Harry J. Bosworth Company), and the skin was sutured where the skull and muscle were exposed. Finally, two stainless steel loops terminating in gold pins were inserted into the anterior and posterior periorbital region of the left eye (and optionally right eye for animals trained in contralateral sequence of CRs, see Chapter 4) for delivery of the stimulation US. Subjects were given postoperative analgesics and antibiotics for 2 days after surgery and were allowed to recover for a week before experiments began.

1.3.2 Conditioning

The subjects were trained in custom-designed, well-ventilated, and sound attenuating chambers measuring 90 × 60 × 60 cm (length, width, height). Each rabbit was placed in a plastic restrainer with their ears stretched over a foam pad and taped down to limit head movement. To measure eyelid position, an infrared emitter/detector system was attached directly to the head stage of each rabbit to record movements of the left external eyelid. These detectors provide a linear readout of eyelid position (± 0.1 mm) at 1kHz sampling rate by measuring the amount of infrared light reflected back to the detector, which increases as the eye closes (Ryan et al., 2006). At the start of each daily

session, the gain of the eyelid position detector was calibrated by delivering a test US to elicit full eyelid closure (defined as 6.0 mm, typical for an adult rabbit). A trial would not start until the rabbit's eyelid was sufficiently open. Stimulus presentation was controlled by custom-designed software for all experiments.

1.3.2.1 Stimulus delivery

Each conditioning chamber was equipped with a speaker that was connected to a stereo equalizer and receiver which were connected to a computer that generated the tone. For subjects trained using tone, the CS was set as a 1 kHz, 500ms, 75 dB sinusoidal tone with a rise and fall time of 5 ms to avoid audible clicks from the speaker. For subjects trained with electrical stimulation of mossy fibers, the CS was a constant frequency pulse train of cathodal current pulses (100 Hz (unless noted otherwise), 500 ms (unless noted otherwise), 0.1 ms pulse width, 100-150 μA), generated by a stimulus isolator (model 2300, A-M Systems, Carlsborg, WA) and passed through the electrode(s) implanted in the middle cerebellar peduncle. The current pulses were controlled by custom written software and delivered through an isolated Pulse Stimulator (model 2100, A-M Systems, Carlsborg, WA). Electrical leads from a separate stimulator (model 2100) were attached to the periorbital electrodes to deliver pulses of electrical stimulation to the left eyelid as the US. The US was a 50 ms train of constant current pulses (50 Hz, 0.7-1 ms pulse width, 1–3 mA) delivered through the periorbital electrodes. US intensity was adjusted

for each rabbit to produce a full eyelid closure without any pain reactions. All types of trials in all sessions were separated by a mean inter-trial interval of 30 ± 10 s.

1.3.2.2 Initial training

For initial training, subjects were given daily eyelid conditioning sessions comprised of 12 blocks of 9 trials each. All subjects were initially trained at an inter-stimulus interval (ISI) of 500 ms. For subjects trained to produce full-sized CRs, each block consisted of 1 CS-alone trial and 8 paired trials. For subjects trained to make half-sized (3 mm) CRs, a block consisted of 1 CS-alone trial and 8 trials which were either a paired or CS-alone trial, depending on the CR amplitude before the US delivery. Following a published training procedure (Kreider and Mauk, 2010) , eyelid position was monitored throughout the presentation of the CS, allowing us to calculate the CR amplitude. If CR amplitude was equal or larger than the target (3 mm) 10 ms before US presentation, the US was omitted on that trial. Otherwise, if the CR size was smaller than the target size, the US was presented at the end of the CS to reinforce an increase in CR amplitude.

1.3.3 In-vivo recordings and unit isolation

The details about recording procedures, single unit isolation and identification of eyelid PCs have been published previously (Halverson et al., 2015). Briefly, each

independently movable tetrode in a microdrive was comprised of four nichrome wires (12 µm diameter; Kanthal Palm Coast). Individual wires were twisted and then heated so that the insulation was partially melted together to form a tetrode. The individual wires of each tetrode were connected to the EIB with gold pins. Each tetrode was gold plated to reduce final impedance to 0.5-1 MΩ measured at 1kHz (nanoZ kit; Neuralynx). During surgery, the tetrodes were placed over the target site of the left anterior lobe of the cerebellar cortex and were advanced to within 2.0 mm ventrally from the target during surgery using stereotaxic guidance. After recovery from surgery each tetrode was lowered in 40-80 µm increments per day until at least one stable single unit was identified, although there were often multiple units on a single tetrode. Typically, tetrodes were allowed to stabilize for 24 h and units were checked again the following day, although on a small fraction of sessions the recording was initiated on the same day if new units appeared to be stable. A custom-written cluster cutting program was used to isolate single units offline. Commonly used waveform features, such as peak, valley and energy were used during cluster cutting. Additional features, such as the late peak measure (Halverson et al., 2015) , were used to identify complex spikes and differentiate them from simple spikes. In some cases, complex spikes formed a separate cluster from simple spikes when viewed in the peak, valley or energy planes. Assignment of putative simple and complex spikes to the same PC was verified by computing a spike-triggered average of simple spikes on complex spikes, demonstrating a post-complex spike pause ((Ohmae and Medina, 2015), Fig. 3.3 C, D). Eyelid Purkinje cells were defined by the

presence of US elicited complex spikes as shown in Fig. 3.4 A. The remaining PCs were considered “non-eyelid”.

1.3.4 Data analysis

Following cluster cutting, all subsequent data analysis was performed using custom-written scripts in MATLAB.

1.3.4.1 Eyelid data analysis

For each trial, 2,500 ms of eyelid position data (200 ms pre-CS, 2,300 ms post CS) were collected at 1kHz sampling rate and at 12 bit resolution. The data were stored to a computer disk for subsequent off-line analysis. Eyelid position data was passed through a low-pass Savitzky–Golay filter. Eyelid velocity was calculated as a derivative of eyelid position with a second-order accurate scheme, passed through a low-pass Savitzky–Golay filter. The small fraction of trials that had eyelid movement during the 200 ms before the CS onset were discarded. For every trial I calculated several measures of the eyelid CR, main ones were: CR amplitude, CR onset time and CR latency to criterion. On paired trials CR amplitude was defined as the maximum value of eyelid position from the baseline before US onset. On CS-alone trials, CR amplitude was defined as the maximum value of eyelid position from the baseline, calculated between CS onset and 500 ms after CS offset. An Eyelid response was counted as a CR if the CR

amplitude reached the 0.5 mm criterion. The CR onset time was defined only for CR trials and was determined using a custom-written two-step algorithm. The first step was designed to detect the initial deflection of eyelid position away from the pre-CS baseline, while the second step used linear interpolation to determine the exact time of CR onset. CR latency to criterion was defined as the first time point when eyelid position deviated above CR criterion.

1.3.4.2 Single unit data analysis

*S*pike times of individual PCs were synchronized with recordings of eyelid position and stimuli onset/offset times. Spike times were rounded to the nearest millisecond. Instantaneous firing rate of each PC was estimated on every trial using the inverse of the inter-spike-interval followed by a two-sided Gaussian kernel with a 20 ms standard deviation window. For every PC the firing rate was normalized by the value of the baseline firing rate during 1500 ms of pre CS activity.

1.3.5 Histology

After the conclusion of experiments the final placement of stimulation electrodes and tetrodes was determined by making small marking lesions. Each implanted electrode was marked by passing 100 μA of anodal DC current for 10 s. A fraction of tetrodes (2-4 out of 16) were marked by passing 10 μA of anodal DC current for 10 s. Animals were killed with an overdose of sodium pentobarbital and perfused intracardially with 0.9%

saline (~1.0 L) followed by 10% formalin (~1.0 L). Heads were post fixed in formalin for at least 3 days after which stimulation electrodes and tetrodes were removed and the brains were extracted. Brains were then cryo-protected in 30% sucrose in formalin for 3 days, embedded in an albumin gelatin mixture, and the cerebellum was sectioned using a freezing microtome at 40 µm. Tissue was mounted on slides and stained with cresyl violet.

CHAPTER 2

Single trial prediction of movement trajectory from Purkinje cells spike trains

2.1 Abstract

One of the fundamental questions in systems neuroscience is the relationship between neuronal activity and behavior. As the principal neurons of the cerebellar cortex, Purkinje cells (PCs) are thought to control the timing properties of motor responses. Eyelid conditioning provides a solely cerebellar-dependent behavior and therefore an opportunity to establish how a PCs population encodes the time profile of conditioned responses (CRs). Here I used single-unit recordings of PCs during the expression of CRs by subjects trained under a variety of behavioral protocols to investigate this relationship. I employed a liner-nonlinear model-based approach to study to what extent eyelid trajectory during a CR on a single trial can be predicted from spike trains of simultaneously recorded PCs. This approach allowed me to systematically test a variety of ways PCs can control the kinematics of CRs. These results show that PCs do not directly encode a single kinematic variable (position or velocity) of a CR. Rather a dynamical model, where PCs provide a 'drive' to the eyelid plant whose dynamics is captured by the second order differential equation, captures the best single trial variability of CR trajectories from all behavioral protocols. Additionally, even a small subset of

PCs, a single cell is some cases, is sufficient to explain a high amount of single trial variability of CR profiles, suggesting that learned changes in PC activity are highly correlated.

2.2 Introduction

One of the fundamental questions in systems neuroscience is the relationship between neuronal activity and behavior. For this reason, Purkinje cells (PCs), the sole output neurons of cerebellar cortex, have been a major target for recordings. A number of studies have addressed the relationship between PCs activity and behavior in a variety of cerebellar-dependent tasks including: smooth pursuit (Medina and Lisberger, 2007, 2009), vestibulo-ocular reflex (VOR) (Raymond et al., 1997), optokinetic response (OKR) (Kitama et al., 1999; Omata et al., 2000), saccade adaptation (Herzfeld et al., 2015) and eyelid conditioning (Halverson et al., 2015; ten Brinke et al., 2015).

While these studies have provided important insights, the analysis they incorporated have been limited by: 1) assumption of a linear relationship between neuronal activity and movement (Medina and Lisberger, 2009; Heiney et al., 2014); 2) averaging data from many trials rather than focusing on explaining trial-to-trial variability; 3) usage of some movements characteristics (amplitude, peak velocity), but not the full trajectory; 4) assumption of a fixed time delay between firing rate and movement (Medina and Lisberger, 2007, 2009); 5) combining results independent of PCs complex spike tuning.

The assumption of a fixed delay (number four in the list) might seem reasonable at the first sight as there is a fixed number of synapses between PCs and motor neurons that produce eyelid CRs. However, what this notion assumes is that PC activity *only* at a fixed time lag (e.g. -10 but not -9 ms) contributes to the behavioral response at a given moment of time. While synaptic integration of PCs activity by downstream targets is not likely to span for hundreds of milliseconds, it is also not instantaneous. The usage of liner filter in models described below makes it possible to study the general form of such integration that corresponds to the data the best without imposed assumptions.

Moreover, such analyses inherently assume that PCs activity encodes a specific kinematic variable (or a linear combination of them in case of the inverse model approach). However a recent study (A. Russo, B. London, S. Perkins, 2016) involving recordings from primary motor and premotor cortices along with muscles EMG suggests a different possibility. In this study two rhesus monkeys were trained to navigate a virtual environment using a hand to pedal either forward or backward. The kinematic patterns of both movement directions were similar but mirrored, while activity of single neurons showed large differences between forward and backward pedaling directions. Recoded EMG activity paralleled neural data and could be successfully decoded from it. These results support the notion that neuronal activity in primary motor cortex does not encode a kinematic variable (velocity) of movement, but rather controls activity of

muscles involved in the movement. It is possible that PC activity during eyelid conditioning also does not merely encode the kinematic variables but rather controls the dynamic evolution of the movement.

I present here the results of a systematic analysis that overcomes all limitations mentioned above. I applied a linear-nonlinear model–based approach (Serruya et al., 2002; Paninski et al., 2004; Hochberg et al., 2006) on data from a published dataset (Halverson et al., 2015). Use of this approach incorporates integration of spike-history instead of a fixed delay assumption, is not limited by a linearity assumption and can be expanded to a broader form than single kinematic variable encoding. It allowed me to systematically test a variety of possibilities for how the whole CR trajectory can be predicted from PC activity on a single-trial basis.

I demonstrate that the best prediction accuracy is achieved by the dynamical model, where PC activity does not specifically encode any single kinematic variable. In the dynamical model PC activity instead provides a drive to the muscle plant (eyelid in our case), the state of which then evolves according to its internal dynamical properties.

2.3 Methods

2.3.1 Purkinje cells recordings dataset

The majority of data analyzed here have been published previously (Halverson et al., 2015) and performed by Hunter Halverson, with some minor additions of recordings performed by me since then under the same training protocols. Tetrode recordings from PCs were made from the left anterior lobe of the cerebellar cortex. Reversible lesion experiments (Garcia et al., 1999) have shown this region of cerebellar cortex to be necessary for acquisition and temporal precision of expression of conditioned responses. PCs were classified as "eyelid" based on the presence of complex spike(s) elicited by the US. Only eyelid PCs were included in the analysis, as only that population received a "teaching signal" through climbing fiber input necessary and sufficient for learning the behavioral responses (Mauk et al., 1986). Therefore, only that population participated in cerebellar learning specific to eyelid conditioning task. Data from only well-isolated single units were used. Recordings of Purkinje cells were collected during delay eyelid conditioning protocols with ISI=250ms, 500ms, 700ms and a dual delay protocol with ISIs=250 and 700ms used interchangeably on even and odd trials of the session. In all sessions the CS was a 1 kHz (or 9.5kHz), 85 dB sinusoidal tone that co-terminated with US. Since the primary focus of the current analysis was on the relationship between PC activity and behavioral eyelid CRs, only sessions with a sufficient number of CRs were used (at least 75 percent of trials with CRs in the session). Along with single unit

recordings, eyelid position was measured (+/-0.1mm) at 1kHz using an infrared emitter/detector system.

2.3.2 Analysis design

All analysis was performed using custom-written scripts in MATLAB. *S*pike times of individual PCs were synchronized with recordings of eyelid position and stimuli onset/offset times. Spike times were rounded to the nearest millisecond.

Our analysis benefited from the fact that several eyelid PCs were usually simultaneously recorded during the session. We therefore have explored how the accuracy of prediction depends on the number of simultaneously used PCs. On a given session with N simultaneously recorded PCs we explored all possible combinations of $m = [1, N]$ PCs. For example, a session with N = 4 recorded PCs would have: 4 combinations of m = 1 PC, 6 combinations of m = 2 PCs, 4 combinations of m = 3 PCs and 1 combination of m = 4 PCs. The analysis described below was repeated for every possible combination of simultaneously recorded PCs.

For every session, 60% of trials were used for the regression fitting procedure to find the linear filter and coefficients corresponding to the specific model (constant term; nonlinearity coefficients; dynamical model coefficients) for the summed contribution from all PCs in the current combination (see Eq. 2.2-2.6). The other 20% of trials were used for regularization of the linear filter size and the remaining 20% of trials — were

used to test the prediction accuracy of the CR time profile, using parameters estimated from the other 80% of the data. These sets of trials were not sequential but spanned the whole session (for example 20% of the trials were composed of every fifth trial in the session). The procedure of assigning trials to the fitting set of trials, regularization set and test set was repeated 5 times so that each trial in the session was used in the test data set. For the results presented in the manuscript I used a ridge regularization method to constrain the size of linear filter. Optimization procedures were performed using the *lsqnonlin* MATLAB function by minimizing the squared error between the concatenated eyelid response profiles from the fit trials data set and concatenated predicted response profile, with the addition of a regularization term.

Spike history vector (see Eq. 2.2) had the same length as the linear filter $\vec{k}$ and was comprised of zeroes and ones, with ones corresponding to milliseconds when PCs spikes had occurred. Linear filter duration equaled 500 ms for the linear position and velocity encoding models; 150 ms for the linear "integrated velocity" encoding model (see Eq. 2.3) and for the dynamical models. The length of the linear filter was chosen to capture the time point when the linear filter value approached zero, indicating that a moment of time sufficiently far apart did not contribute to the model prediction. In order to avoid over-fitting I parameterized each linear filter with a set of 15 raised cosines temporal basis functions (Fig. 2.1). Such parameterization was chosen so that the filters

can have a fine temporal structure near the time of the prediction (t = 0) while also maintaining a restricted number of parameters.

Thus a linear filter $\vec{f}$ was represented as a linear combination of basis functions:

$$\vec{f} = \sum_{i=1}^{N} w_i \cdot \vec{b}_i \tag{2.1}$$

Where $\vec{b}_i$ are raised cosines basis functions and w_i are corresponding weights. During the fitting procedure finding appropriate values of weights was sufficient to capture the shape of the linear filter. All weights were initialized at random values within a specified [-0.5, 0.5] range.

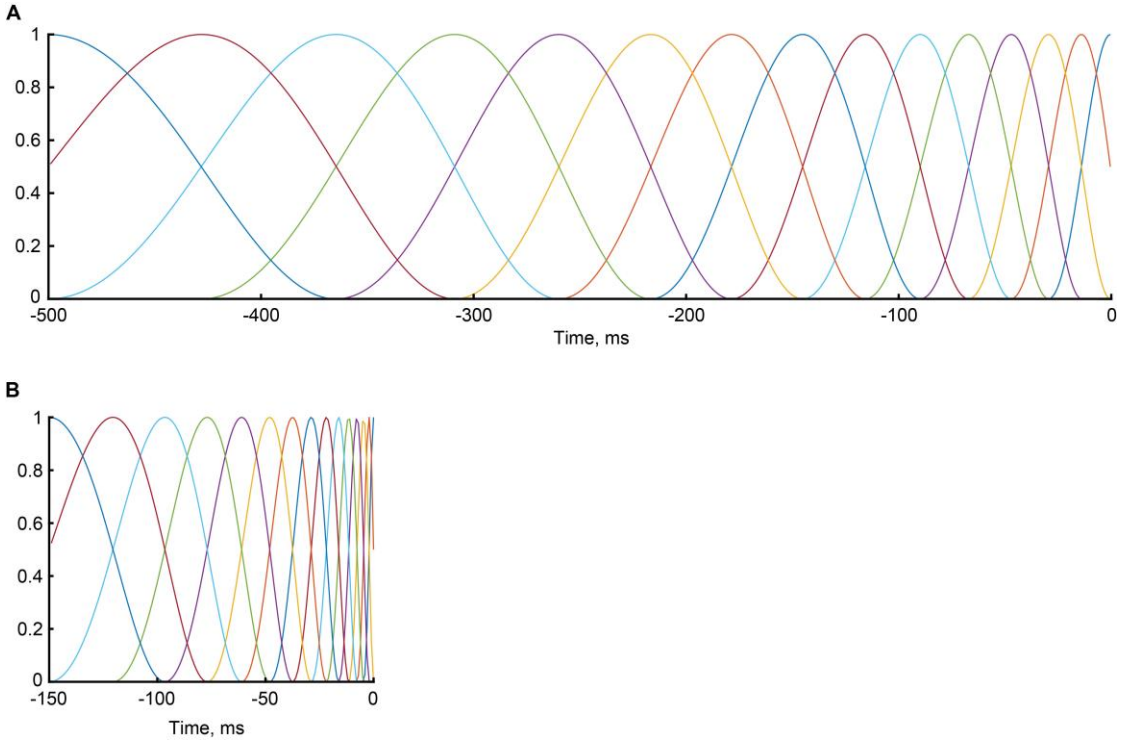

Fig. 2.1. Temporal basis functions.
A) Raised cosine basis functions used for linear position and velocity encoding models B) Basis functions used for linear "integrated velocity" encoding model (see Eq. 2.3) and for dynamical models.

Eyelid position data and corresponding recordings were analyzed starting from 200 ms of baseline before CS onset until US onset time (since the US evokes non-cerebellar reflexive eyelid closure). Therefore, on paired trials we used 200 ms + ISI length of eyelid position profiles. On CS-alone trials we used an additional 400 ms after CS offset to capture the full CR profile. The accuracy of each prediction was measured as a coefficient of determination between the actual and predicted eyelid position time profiles, concatenated across all test trials.

2.3.3 Linear models

The linear models were constructed in a following way:

$$y(t) = \sum_{PC=1}^{N} w_{pc} \cdot \overrightarrow{spikes}_{pc}(t) \cdot \vec{f} + c \qquad (2.2)$$

Where y(t) is the value of the kinematic variable (eyelid position or velocity) at moment t, $\overrightarrow{spikes}$ is a PC spike history vector before the moment t, $\vec{f}$ – is the linear filter, w_{pc} is a weight associated with a particular PC, c is a constant term and the sum over PC indicates sum over N eyelid PCs from those recorded simultaneously during the session (see also Fig. 2.4 C-E). Implemented this way, the linear filter $\vec{f}$ was the same across all PCs in the combination, but their relative contributions were adjusted by the w_{pc} terms (defined at (0, 1] range).

I considered two main types of linear models: linear position encoding and linear velocity encoding. However, velocity encoding can be implemented in two ways: 1) a conventional way where squared error between real and predicted eyelid velocity is minimized; 2) the velocity prediction is integrated into eyelid position and squared error between real and predicted, in this way eyelid position is minimized. Equation 2.2 illustrates the description:

$$pos(t) = \int_0^t \left(\sum_{PC=1}^{N} w_{pc} \cdot \overrightarrow{spikes}_{pc}(\tau) \cdot \vec{f} + c \right) d\tau \qquad (2.3)$$

Corresponding to the implementation of the fitting procedure, prediction accuracy for the conventional velocity encoding model was calculated by comparing the prediction with real eyelid velocity profile, while prediction of "integrated velocity" model was compared to the real eyelid position profile.

The "integrated velocity" implementation consistently gave superior results compared to the conventional velocity encoding model (Fig. 2.2 D, $p < 0.001$ for PCs number ≤ 5, $p < 0.01$ for PCs number = 6, t-test with Bonferroni correction). Therefore, though the results section I refer to the model described by Eq. 2.3 as a linear velocity encoding model. To indicate that during the fitting procedure the squared error was minimized with eyelid position, a subscript 'p' was used as 'vel_p' in figures.

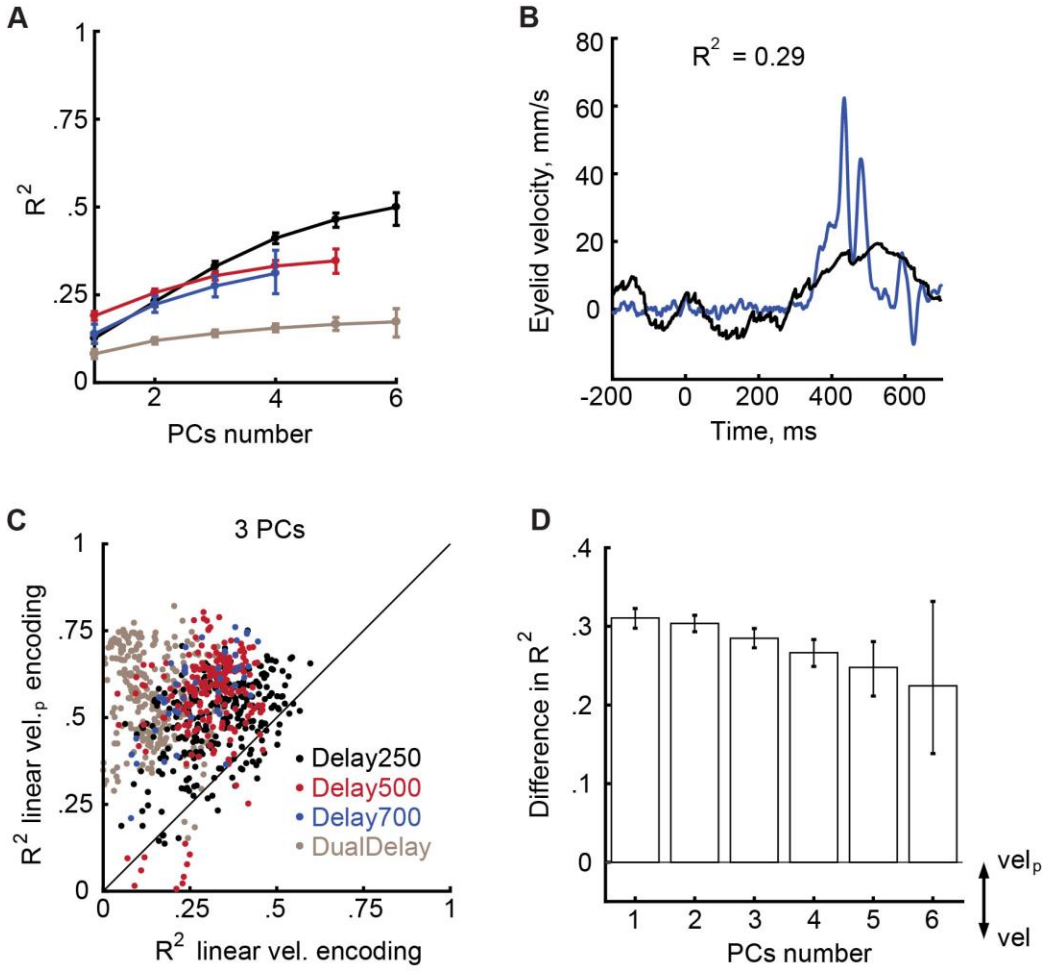

Fig. 2.2. Linear velocity encoding models results.
A) Results of conventional linear position encoding model. Average prediction accuracy as a function of number of PCs used for prediction. Lines are color-coded according to the behavioral protocol, as indicated in the legend. Error bars indicate 95% confidence intervals. B) Example trial illustrating real (blue line, Delay700) and predicted (black line) profiles of eyelid velocity. C) Comparison of prediction accuracy between two velocity encoding models. The plot shows results from all combination of 3 PCs, each dot corresponding to a single combination. Dots are color-coded by corresponding behavioral protocol; the diagonal line is shown in black. D) Average difference in prediction accuracy between two velocity encoding models as a function of PCs number used for prediction. Positive numbers indicate higher prediction accuracy by "integrated velocity" model where error with eyelid position is minimized. Error bars indicate 95% confidence intervals.

2.3.4 Nonlinear models

Nonlinearity was introduced by modifying the linear model in the following way:

$$y(t) = F\left(\sum_{PC=1}^{N} w_{pc} \cdot \overrightarrow{spikes}_{PC}(t) \cdot \vec{f} + c\right) \quad (2.4)$$

Here F is a nonlinear transformation function. In order to determine which nonlinearity to use, I calculated the shape of empirical nonlinearities comparing results of linear models with real data (see Fig. 2.7). Results presented in the manuscript were obtained using a sigmoidal nonlinearity:

$$Sigm(x) = \frac{a}{1 + e^{-x+b}} \quad (2.5)$$

I have also considered the soft-threshold nonlinearity, which produced very similar, but less robust results (data not shown).

2.3.5 Dynamical model

In the dynamical model, PCs do not explicitly encode any kinematic parameter in their activity, but rather provide a drive to the eyelid plant which dynamically modulates the movement trajectory. A simple way to formalize this notion is through the second order linear differential equation, analogous to Newton's second law:

$$\ddot{y}(t) + a \cdot \dot{y}(t) + b \cdot y(t) = f\left(\sum_{PC=1}^{N} \overrightarrow{spikes}_{PC}(t) \cdot \vec{f} + c\right) \quad (2.6)$$

Here the input from PCs comes as an external force in the right side of the equation, while coefficient *a* and *b* are determined by the properties of the eyelid plant and therefore influence the dynamics of eyelid closure given the PCs input. Number of dots above *y* indicates order of the derivative. I considered two simple forms of the first and second order linear differential equations, formalized in Eq. 2.6 and Eq. 2.7:

$$\dot{y}(t) + a \cdot y(t) = F\left(\sum_{PC=1}^{N} \overrightarrow{spikes}_{PC}(t) \cdot \vec{f} + c\right) \quad (2.7)$$

I have also explored more complex options, such as cross-term $\dot{y}(t) \cdot y(t)$, quadratic or cubic terms of $\dot{y}(t)$ and/or $y(t)$, but did not see a significant positive change in model performance.

2.4 Results

2.4.1 Benefits of behavioral variability

The majority of the data analyzed in this chapter have been published previously (Halverson et al., 2015). I analyzed the relationship between PC activity and behavior for 140 eyelid PCs recorded from N = 6 subjects during delay eyelid conditioning sessions. In the present analysis I used data from four different behavioral protocols: Delay 250 (N = 3, 36 eyelid PCs), Delay 500 (N = 5, 54 eyelid PCs), Delay 700 (N = 1, 12 eyelid PCs) and Dual Peak (N = 2, 38 eyelid PCs), each resulting in different CR kinematics. The number in the name of each protocol refers to the ISI used in milliseconds. A simple change of the time interval between the onsets of CS and US (inter-stimulus interval, ISI) results in different kinematics of a learned eyelid response (Mauk and Ruiz, 1992). Figure 2.2 B demonstrates average time-profiles of eyelid position and velocity for all four protocols color-coded as indicated in the legend. Eyelid position and velocity trajectories were truncated at US onset time. Such simple presentation of behavioral data illustrates several important similarities and differences between CR profiles trained to different protocols. The peak of eyelid position was appropriately timed for each ISI subjects were trained with. The velocity peak value is larger for small ISIs (compare Delay250 to Delay500 to Delay 700) since CRs have to happen in a shorter time window. Subjects also tend to initiate their CR later (relative to CS onset) if trained at longer ISIs. In addition, if the US occurred at two distinct times on different trials, the subjects

acquired CRs with two correspondingly timed peaks (Millenson et al., 1977; Freeman et al., 2003; Halverson et al., 2015) (Figure 2.2 B tan line). Importantly, eyelid velocity had a robust negative component in this protocol. Increasing the length of the ISI makes CRs more variable, even in well-trained animals (Khilkevich et al., 2016). I quantified this effect for the two most commonly used features of a CR – onset time and amplitude. The distributions of CR amplitudes and CR onset times are shown on Fig. 2.3 C. Here behavioral data was combined from all sessions with a given protocol. The variance of each distribution (Fig. 2.3 D, E), demonstrated that CRs at longer ISIs (Delay 500 and Delay 700) have high variability. In sum, each behavioral protocol poses a specific constrain on a possible relationship between activity of PCs and conditioned eyelid responses. At short ISIs, CRs have to be very rapid (Fig 2.3 B, ISI250), but CRs on different trials do not show a lot of variance. Longer ISIs introduce a challenge of high trial-to-trial variability of CR profiles that needs to be captured by the model prediction. Eyelid CRs during the Dual peak protocol have both positive and negative velocity at different times (though CRs on CS-alone trials trained to a single ISI also have a negative velocity component (see Methods)). Richness of constrains from different behavioral protocols puts a true test on any hypothesis attempting to describe how PCs activity controls CR kinematics.

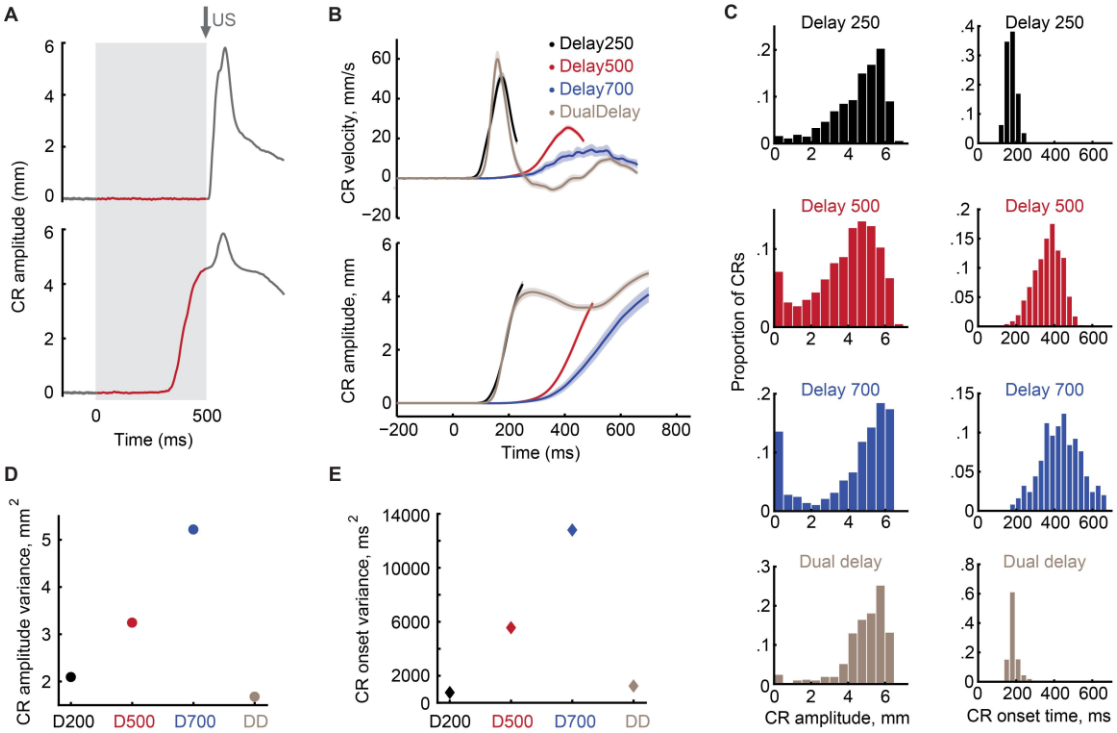

Fig. 2.3. Differences in CR expression at different behavioral protocols.
A) Eyelid position as a function of time before (upper panel) and after (lower panel) acquisition of CRs. Upward deflection corresponds to eyelid closure. CS duration is indicated by grey-shaded area. Eyelid position during CS is shown in red. Eyelid position before CS onset and after US onset is shown in dark grey to indicate the absence of cerebellar contribution. B) Average profiles of eyelid position (upper panel) and velocity (lower panel) for four behavioral protocols used. Shaded regions indicate 95% confidence intervals. Color-coding of behavioral protocols is indicated in the legend. C) Distribution of CR amplitudes (left panels) and CR onset times (right panels) for all behavioral protocols. D) Variance of CR amplitude distributions as a function of behavioral protocol. E) Similar plot for variance of CR onset distributions.

2.4.2 Linear models design

Previous studies have established that activity of eyelid PCs is highly correlated with behavior (Halverson et al., 2015; ten Brinke et al., 2015). Unpublished results from

our lab suggest that learned changes in activity of PCs receiving similar climbing fiber input are also highly correlated. Both of these ideas can be observed in raw data. Figure 2.4 A shows activity of two simultaneously recorded eyelid PCs along with subject's behavioral responses. Here trials were sorted in chronological order from bottom to top, onset times of behavioral responses are indicated by red dot in PC raster plots. Figure 2.4 B shows the same data, but with trials sorted by the CR onset times. Notice how closely the decrease in PC activity matched the precise timing of CRs on a given trial. Based on such data, the goal of this chapter was to find the transformation of eyelid PCs activity to behavior that will predict single trial CR trajectories for all behavioral protocols.

Previously the relationship between PC activity and behavior has been addressed by correlation-based analysis (Halverson et al., 2015; ten Brinke et al., 2015) or other linear techniques such as the inverse model approach (Medina and Lisberger, 2007, 2009; Raghavan and Lisberger, 2017). Apart from the assumption of linearity, these studies assumed a fixed time-lag between eyelid PCs activity and behavior. I therefore started with the construction of a linear model that does not possess this assumption. The model was constructed in the following way:

$$y(t) = \sum_{PC=1}^{N} \overrightarrow{spikes_p}(t) \cdot \vec{f} + c \qquad (2.2)$$

where y(t) is the value of the kinematic variable (eyelid position for example) at moment t, $\overrightarrow{spikes}$ is a spike history vector before the moment t, $\vec{f}$ – is the linear filter, c is a

constant term and the sum over PC indicates the sum over N eyelid PCs from those recorded simultaneously during the session. This is further illustrated in Figure 2.4 C-E, demonstrating the process of eyelid position prediction from the PC spike train on a single trial at different moments of time. Spike times of the eyelid PC are shown as a series of black (or blue) dots, trajectory of the real eyelid CR is shown in red. Here the prediction using a linear position encoding assumption is shown. Figure 2.4 C shows the prediction at t = 190 ms from CS onset. Spikes of PC from the spikes history window before this moment are shown in blue. Predicted eyelid trajectory at previous moments of time is shown in black, the estimate at t = 190 ms is shown in blue. A similar schematic is used for panels D and E, illustrating a gradual process of CR trajectory prediction.

This model design incorporates a linear filter $\vec{f}$ instead of the commonly used fixed time-lag (Medina and Lisberger, 2009; Heiney et al., 2014; Halverson et al., 2015), and enables a more complete representation of the influence of past spikes on the current value of kinematic variable. An inspiration for this approach came from studies using the generalized linear model (GLM) approach (Truccolo et al., 2004; Pillow et al., 2008) to study the relationship between stimulus and spiking activity in sensory systems (also see (Park et al., 2014) as an extension of such approach). In those paradigms the stimulus filter accomplishes a similar role to the linear filter in our case. The shape of the linear filter corresponds to the shape of PC activity (or sensory stimulus in the case of GLM)

that would produce the maximum value of the encoded kinematic variable (highest spike probability for GLM). The difference between the GLM approach and the one employed here is the absence of post-spike filter in our case. Indeed, the GLM is typically applied to model spike train given the presented stimulus. Post-spike filter there is used to model firing properties of the specific neuron, such as burstiness, adaptation or refractory period(Pillow et al., 2008). Since in this chapter the goal was to predict a motor response from the spike train rather than predict spike train from stimulus, post-spike filter was not used.

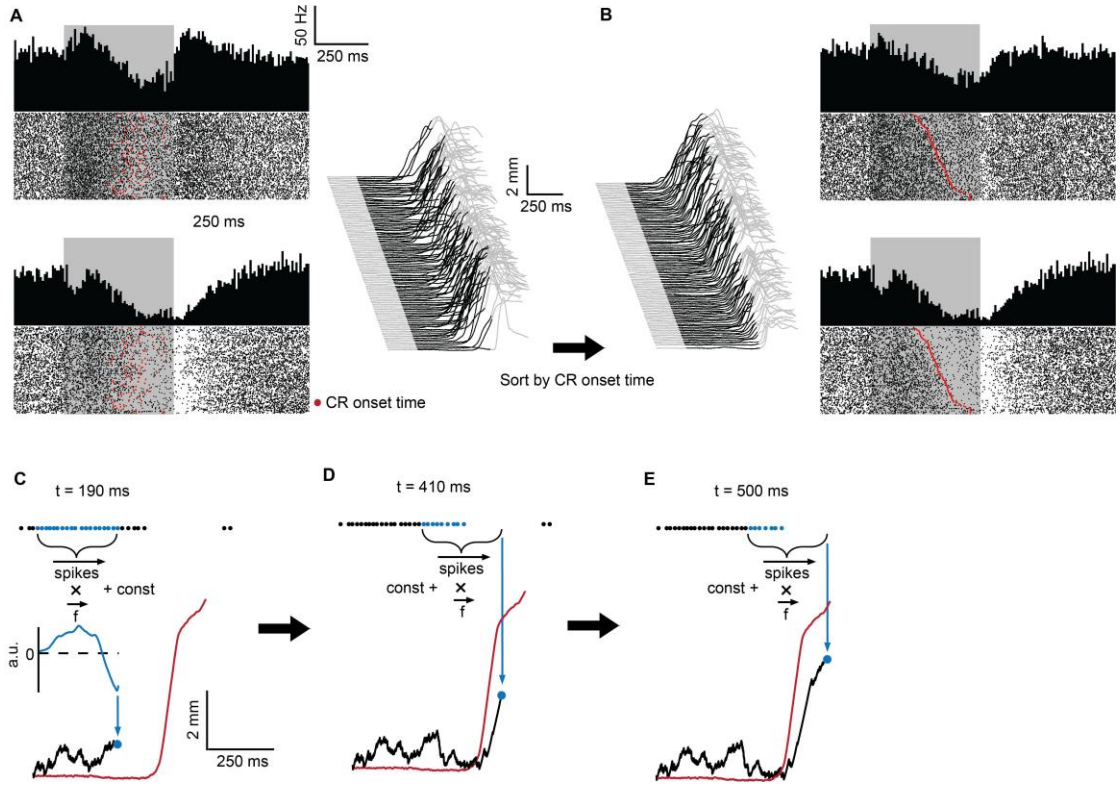

Fig. 2.4. Design of linear models.
A) PSTHs and raster plots of eyelid PC activity along with eyelid responses during a Delay 500 session. Duration of CS is shown as grey-shaded area in PSTHs and as black color in behavioral waterfall plot. Trials are sorted chronologically from bottom to top. CR onset times are shown as red dots is raster plots. B) Data from the same session, but with trials sorted by behavioral CR onset times (red dots). C-E) The process of eyelid position prediction by a linear model. A spike-train from PC on a single trial is shown by black (and blue) dots on top. Real trajectory of eyelid CR on that trial is shown in red. A prediction from linear position encoding model is shown in black. Panel C) illustrates the prediction at t = 190 ms from CS onset. Spikes of PC contributing to the prediction at this moment are shown in blue. Predicted eyelid trajectory at previous moments of time is shown in black, the estimate at t = 190 ms is shown as a blue dot. Similar schematics are used for panels D and E, illustrating a gradual process of CR trajectory prediction.

After the fitting procedure (see methods), I tested the ability of the model to predict CR profiles on a separate set of test trials. All results presented in the manuscript are based on the test datasets, unless indicated otherwise. The analysis benefited from the fact that several eyelid PCs were usually recorded simultaneously during the session (Delay 250 - up to 6 simultaneous eyelid PCs, Delay 500 - up to 5 simultaneous eyelid PCs, Delay 700 - up to 4 simultaneous eyelid PCs and Dual Peak - up to 6 simultaneous eyelid PCs). Therefore I explored each model's prediction accuracy as a function of PCs number used for the prediction.

Figure 2.5 shows the summary of results for the linear eyelid position or velocity encoding models. In both models, a fitting procedure was implemented to minimize the squared error between actual and predicted eyelid position profiles (see Methods). Panels A and B of Fig. 2.5 show the average R^2 value of each prediction as a function of the PCs number used for each prediction. Not surprisingly, for both models the prediction accuracy improved across all behavioral protocols (indicated by the same color-coding as in Fig. 2.3) with the number of PCs used. A direct comparison between the different model performances for all possible combinations of 3 simultaneously recorded PCs is shown on Fig. 2.5 C and Fig. 2.6. Here each dot corresponds to a single combination, demonstrating the R^2 value achieved by velocity versus position encoding models given the same input data. The majority of dots were above the diagonal line, indicating that prediction accuracy of the linear velocity encoding model was generally higher. The difference between models performance as a function of PCs number is shown in Fig. 2.5

D. Here results were averaged through all behavioral protocols. Overall, the velocity encoding model gave consistently better or equal results relative to the position encoding model ($p < 0.001$ for PCs number ≤ 5, t-test with Bonferroni correction), though the difference decreased with more PCs being used for the prediction.

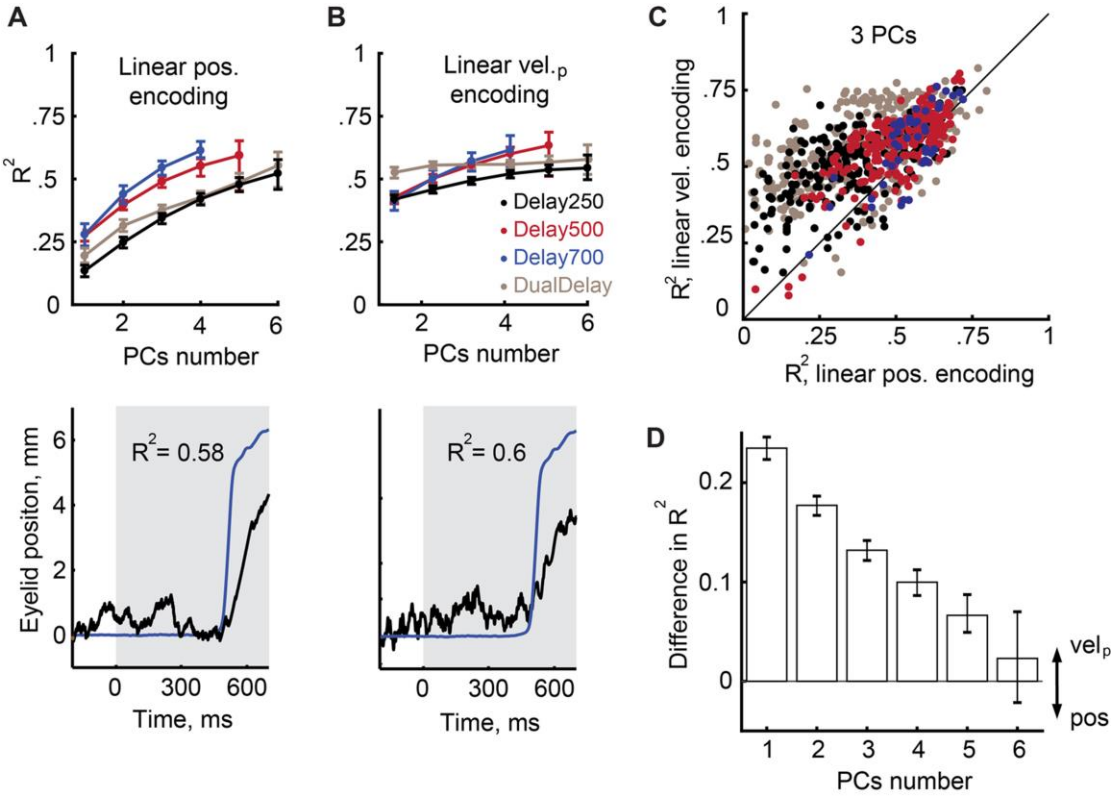

Fig. 2.5. Linear models performance.
A) Results of linear position encoding model. Upper panel: average prediction accuracy as a function of number of PCs used for prediction. Lines are color-coded according to the behavioral protocol, as indicated in the legend. Error bars indicate 95% confidence intervals. Lower panel: example trial illustrating real (blue line, Delay700) and predicted (black line) profiles of eyelid position. For the example prediction trial combination of 3 PCs was used. B) Results of linear velocity encoding model. Design and arrangement of panels is the same as in A). C) Comparison of prediction accuracy between two models. The plot shows results from all combination of 3 PCs, each dot corresponding to a single combination. Dots are color-coded by corresponding behavioral protocol, the diagonal line is shown in black. D) Average difference in prediction accuracy between two models as a function of number of PCs used for prediction. Positive numbers indicate higher prediction accuracy by velocity encoding model. Error bars indicate 95% confidence intervals.

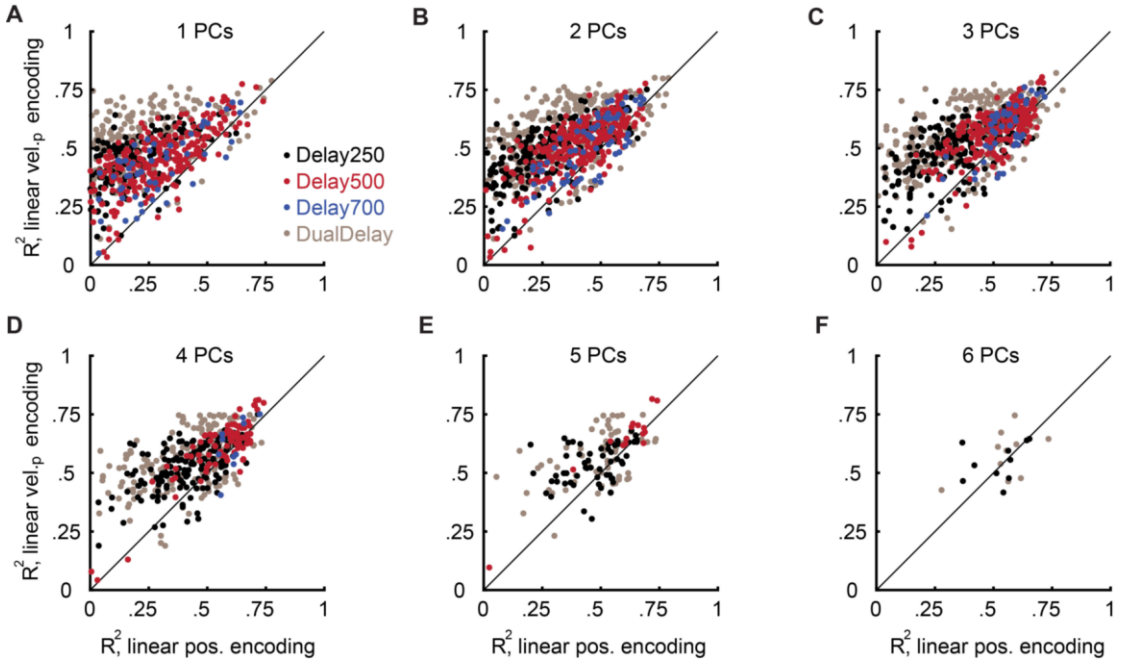

Fig. 2.6. Comparison between linear models within each combination of PCs.
A-F) Plots show amount of variance of eyelid responses profiles captured by linear velocity (y axis) or position (x axis) encoding models. Each panel correspondingly shows results from all possible combinations of one to six eyelid PCs. Each dot corresponds to a single combination of PCs, color-coded corresponding to behavioral protocol. Diagonal line is shown in black.

2.4.3 Deviations from linear encoding

While the linear velocity encoding model demonstrated quite high prediction accuracy (R^2 = 0.59 ± 0.02 for maximum number of simultan9eous PCs per each protocol), example traces of predicted eyelid position illustrate two shortcomings (lower panels of Figure 2.5 A, B). First, because the model is linear, it was unable to keep the baseline eyelid position at zero, reacting to any deviation in PCs firing rate from the

baseline value. Second, during a CR, the linear model prediction tended to undershoot compared to the real eyelid position.

To evaluate whether these deviations are consistent across all data, I calculated empirical nonlinearities. Figure 2.7 A shows average empirical position nonlinearities for the linear velocity encoding model. Different columns show results obtained for different numbers of simultaneously recorded PCs, different rows correspond to different behavioral protocols, color-coded as above. Empirical position nonlinearities had sigmoid-like shapes, were consistent through all behavioral protocols and did not significantly change with the number of PCs used for prediction. Consistent deviations from the diagonal line again demonstrated two features that were noticeable from example traces in the lower panels of Figure 2.5. First, the prediction from the linear model included negative values of eyelid position. Second, there was a consistent undershoot to the high values of eyelid position that did not go away or significantly improve with the number of PCs used for the prediction. Empirical nonlinearities for the linear position encoding model demonstrated similar shapes (not shown). Together these results point towards a non-linear relationship between spiking activity of eyelid PCs and CR profile.

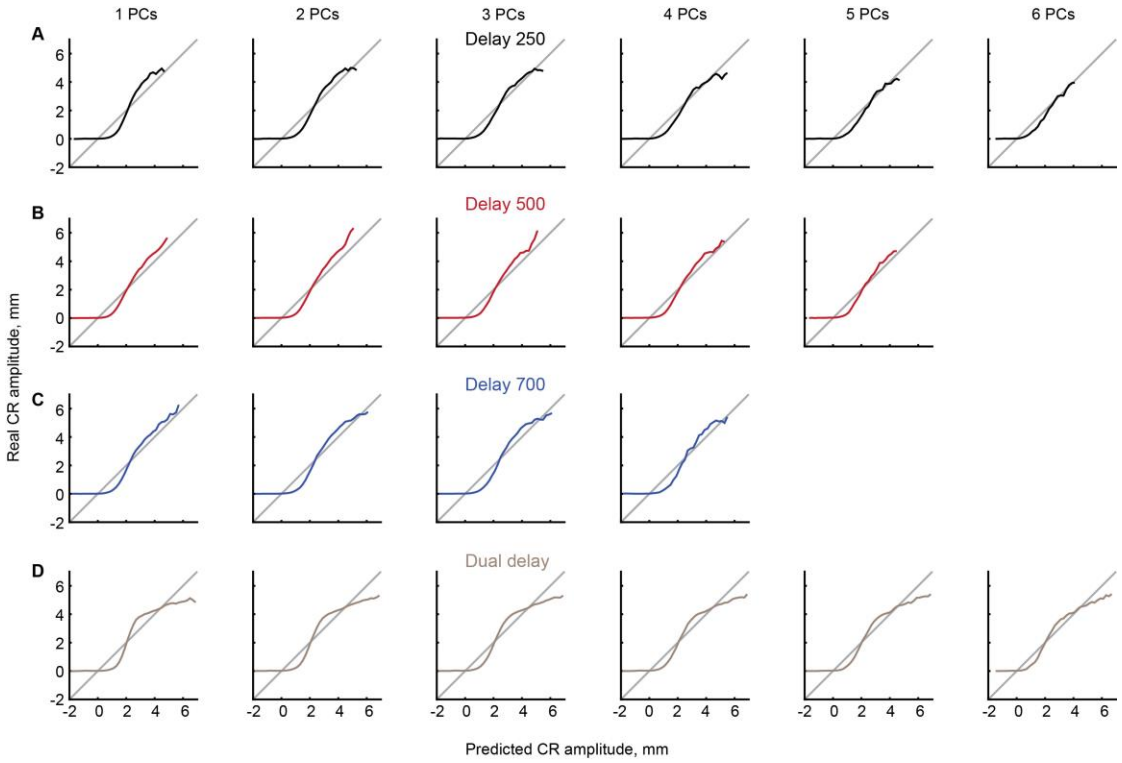

Fig. 2.7. Empirical position nonlinearities.
A) Results are shown for Delay 250 protocol. Each column displays results for N PCs used in linear velocity encoding model for prediction with N indicated at the top of corresponding column. Each panel shows real versus predicted eyelid position, averaged across all trials and combinations of PCs corresponding to the panel. Diagonal line is shown in grey.
B-D) Same as A), but for Delay 500, Delay 700 and Dual Delay protocol respectively.

2.4.4 Dynamical model approach

In order to cope with problems of linear models, I propose a new approach. Rather than encoding a specific kinematic variable of a CR, let's assume that PCs simply provide a 'drive' through downstream areas to the eyelid plant. In this general case, the evolution of eyelid position as a function of time would depend on dynamical properties of the plant. Assuming that the dynamics of the eyelid plant can be captured by a second order linear differential equation, such notion can be formalized as:

$$\ddot{y}(t) + a \cdot \dot{y}(t) + b \cdot y(t) = Sigm\left(\sum_{PC=1}^{N} \overrightarrow{spikes}_{PC}(t) \cdot \vec{f} + c\right) \quad (2.7)$$

where $y(t),\ \dot{y}(t)\ and\ \ddot{y}(t)$ are eyelid position, velocity and acceleration respectively at the moment of time t, $Sigm$ denotes a sigmoid nonlinearity and other notations are similar to equations described above. I will refer to this approach as the "dynamical model" throughout the rest of the manuscript.

The form of Eq. 2.7 resembles the Newton's second law. Here the input from PCs comes as an external force, while coefficients *a* and *b* are determined by the properties of the eyelid plant and therefore influence the dynamics of eyelid closure given PC input. The exact form of the left side of Eq. 2.7 had to be estimated to a degree. Support for the sufficiency of a linear differential equation to describe the dynamics of the eyelid plant comes from previous work examining this topic (Lepora et al., 2007, 2009). There authors studied the relationship between the retractor bulbi muscle EMG

activity and CR profiles produce by a nictitating membrane of rabbit's eye. They found that a first order linear differential equation along with a sigmoidal nonlinearity was sufficient to capture eyelid position profiles. Therefore, I mainly explored two possibilities: either the simplest form of the first order (eq. 2.6) or second order (eq. 2.7) linear differential equation. I have also explored an addition of more complex terms in Eq. 2.7, such as cross-term $\dot{y}(t) \cdot y(t)$, quadratic or cubic terms of $\dot{y}(t)$ and/or $y(t)$, but did not see a significant positive change in model performance.

Comparison in prediction accuracy between the dynamical models described by the first or second order linear differential equation is shown in Fig. 2.8. The second order differential equation demonstrated a small but consistent increase in the prediction accuracy of the CR profiles that remained across all numbers of simultaneous combinations of PCs. Therefore, I used Eq. 2.7 for the dynamical model approach through the rest of the manuscript.

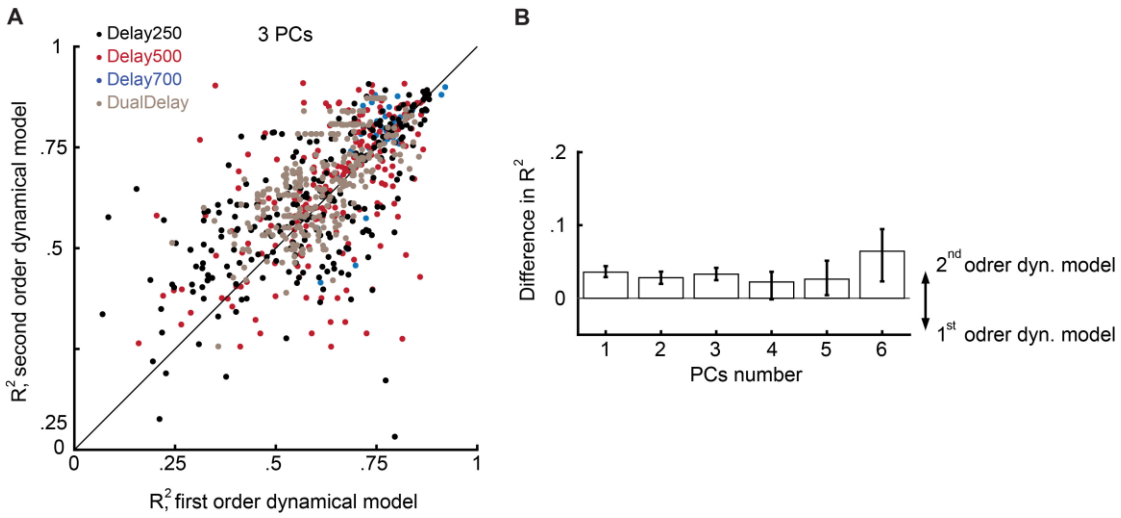

Fig. 2.8. Comparison between first and second order dynamical models.
A) Results are shown from all possible combination of 3 PCs, each dot corresponds to a single combination. Dots are color-coded by corresponding behavioral protocol, the diagonal line is shown in black. B) Average difference in prediction accuracy between two models as a function of number of PCs used for prediction. Positive numbers indicate higher prediction accuracy by second order dynamical model. Error bars indicate 95% confidence intervals.

2.4.5 Dynamical model results

Since a commonly accepted notion in the literature is that PCs linearly encode velocity, I compared the prediction accuracy of the linear velocity encoding model with the dynamical model. The summary of these results is shown in Fig. 2.9. As before, the accuracy of the prediction improved with more simultaneously recorded eyelid PCs used (Fig. 2.9 A), but achieved higher values relative to other models. This point is illustrated in Fig 2.9 C and D.

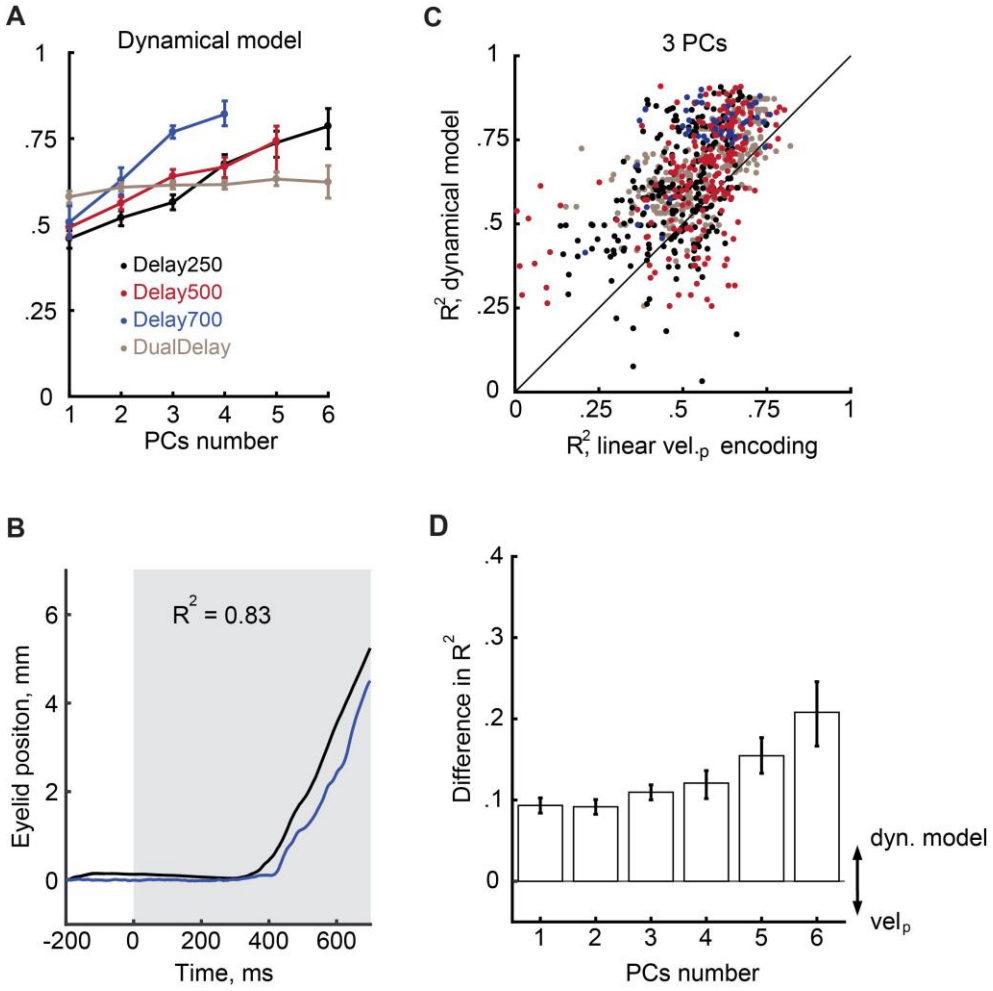

Fig. 2.9. Dynamical model performance.
A) Average prediction accuracy of dynamical model as a function of number of PCs used for prediction. Lines are color-coded according to the behavioral protocol, as indicated in the legend. Error bars indicate 95% confidence intervals. B) Example trial illustrating real (blue line, Delay700) and predicted (black line) profiles of eyelid position. For the example trial combination of 3 PCs was used. C) Comparison of prediction accuracy between dynamical model and linear velocity encoding model. The plot shows results from all combination of 3 PCs, each dot corresponding to a single combination. Dots are color-coded by corresponding behavioral protocol, the diagonal line is shown in black. D) Average difference in prediction accuracy between two models as a function of number of PCs used for prediction. Positive numbers indicate higher prediction accuracy by dynamical model. Error bars indicate 95% confidence intervals.

Figure 2.9 C demonstrates prediction accuracy for the dynamical versus linear velocity encoding models for all combinations of 3 simultaneously recorded PCs. Data for all behavioral protocols is color-coded using the same scheme as before. This panel clearly demonstrates that given the same PC spike trains, the dynamical model was able to achieve better prediction accuracy of CR profiles. The summary of difference in prediction accuracy between the two models as a function of the number of eyelid PCs used for the prediction is shown in Fig. 2.9 D. The dynamical model approach not only outperformed the linear velocity encoding assumption ($p < 0.001$ for all PCs numbers used, t-test with Bonferroni correction), but the difference increased with more PCs used (one-way ANOVA, $p < 1 \cdot 10^{-7}$, $F_{4,\ 3044} = 9.28$). Therefore, after exploring a variety of possible ways PC activity can relate to CR kinematics, I conclude that the dynamical model approach achieved the best results.

The main goal of this project was to find the best way to capture most of the single trial variability of CRs from the activity of PCs. While Fig. 2.9 demonstrates that the dynamical model approach captures a large portion of the variance in CR profiles, I wanted to test how well this approach makes predictions of CR features on trial by trial basis. Two most commonly reported features of eyelid CRs are CR onset time and CR amplitude. Since both variables showed larger amounts of natural variability for behavioral protocols with longer ISIs (Fig. 2.3 C-E), I studied how well the dynamical model predicts those CR features on sessions with Delay 500 and Delay 700 protocols. Notice that the model was fit to maximize the amount of variance of the whole CR profile

captured, rather than specific points that determine CR amplitude or the time of CR onset. Thus, I would argue that if the model also captures these CR features well, it would imply that: 1) this information is present in PCs spike trains; 2) the dynamical model approach resembles the true relationship between PCs activity and eyelid CRs.

Figure 2.10 shows real versus predicted CR amplitude achieved by the dynamical and linear velocity encoding models. Each dot corresponds to a single trial; columns are organized by the number of PCs used for the prediction, increasing from left to right. The corresponding concordance correlation coefficients between real and predicted CR amplitude are indicated separately on each panel. Ideally, all dots would be on the diagonal line shown in black. Scatter plots clearly demonstrate that for low numbers of PCs both models had predominantly two types of errors: 1) overestimate CR amplitude on nonCR trials and 2) underestimate amplitude on trials with large CR amplitudes. Deviations from the diagonal line were clearly smaller for the results from the dynamical model. Since these challenges put opposite requirements on predictions it is likely that these errors were the results of the models trying to minimize the overall error. Both models showed an improvement in prediction accuracy of CR amplitude with more PCs used. Consistent with the main results shown in Fig. 2.9, the dynamical model made single trial prediction of amplitude more accurately for both Delay 500 and Delay 700 protocols.

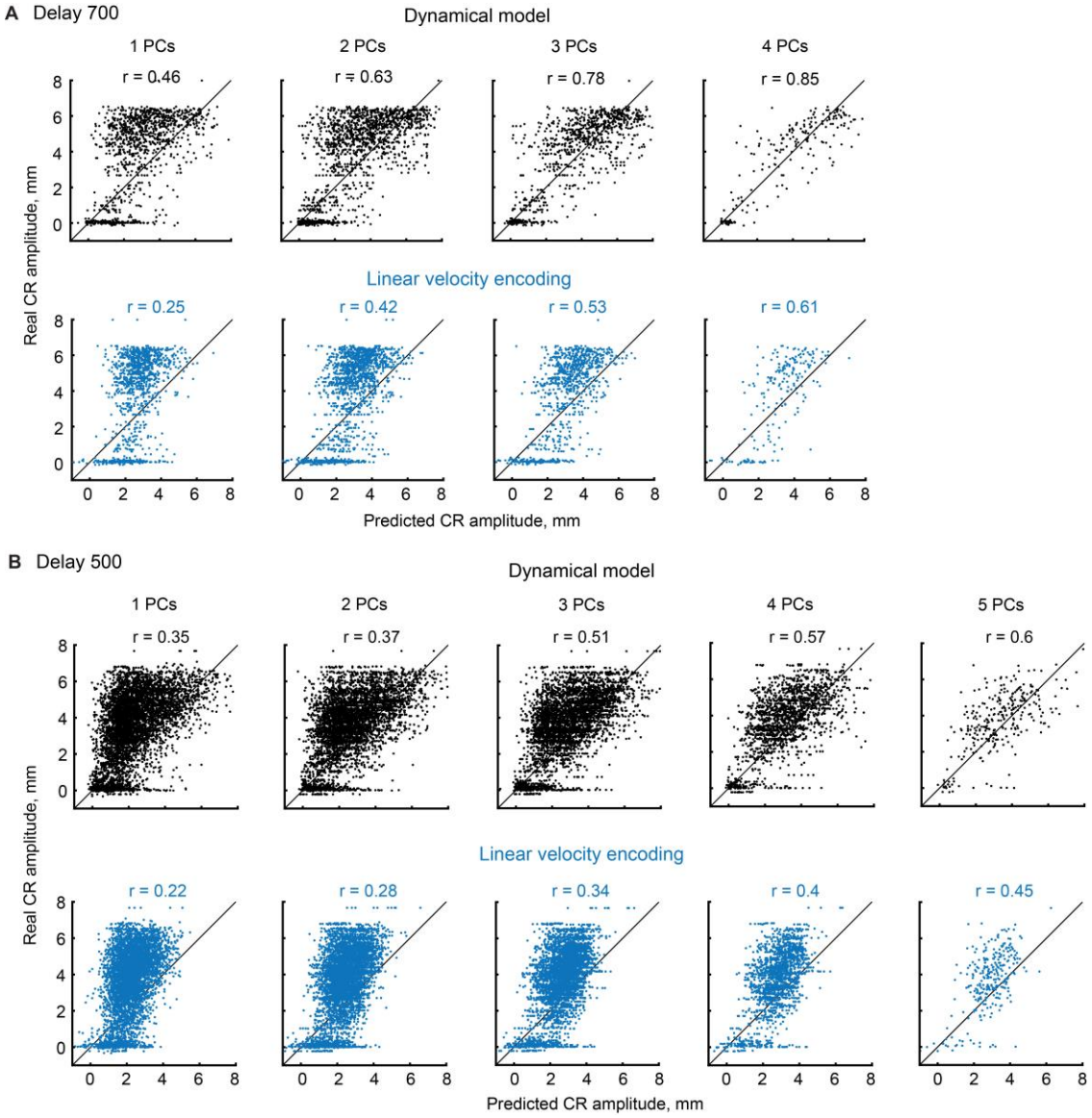

Fig. 2.10. Single trial prediction of CR amplitudes.
A) Performance of a dynamical model (upper row, black color) and linear velocity encoding model (bottom row, blue color) for Delay 700 protocol. Panels are rearranged in columns by the number of eyelid PCs used for prediction. Each dot represents a single trial with real versus predicted CR amplitudes plotted against each other. Corresponding concordance correlation coefficient is indicated for each panel. Diagonal is shown by a black line. B) The same for Delay 500 data.

In addition to CR amplitude, I also explored if the dynamical model can more accurately predict CR onset times relative to the linear encoding model. Figure 2.11 shows these results with the layout similar to Fig. 2.10. Here I used the latency to CR criterion equal to 1 mm to estimate CR timing instead of the exact CR onset time. This was done to minimize the influence of small fluctuations in the CR profile and provide cleaner prediction results. Here the main challenge for both models was to maintain eyelid position below CR criterion until an actual CR. This is evident by a cloud of dots on the left of the diagonal line. Notice that the inability to stably maintain CR baseline was more prominent for the velocity encoding model than the dynamical model for all numbers of PCs used. Interestingly, while the dynamical model showed a consistent increase in prediction accuracy with more PCs, the gain for the linear velocity encoding model was more subtle with concordance correlation coefficients based on the same number of PCs for two models showing tenfold differences. In sum, the dynamical model approach was able to accurately predict both CR amplitude and CR latency to criterion on a single trial resolution.

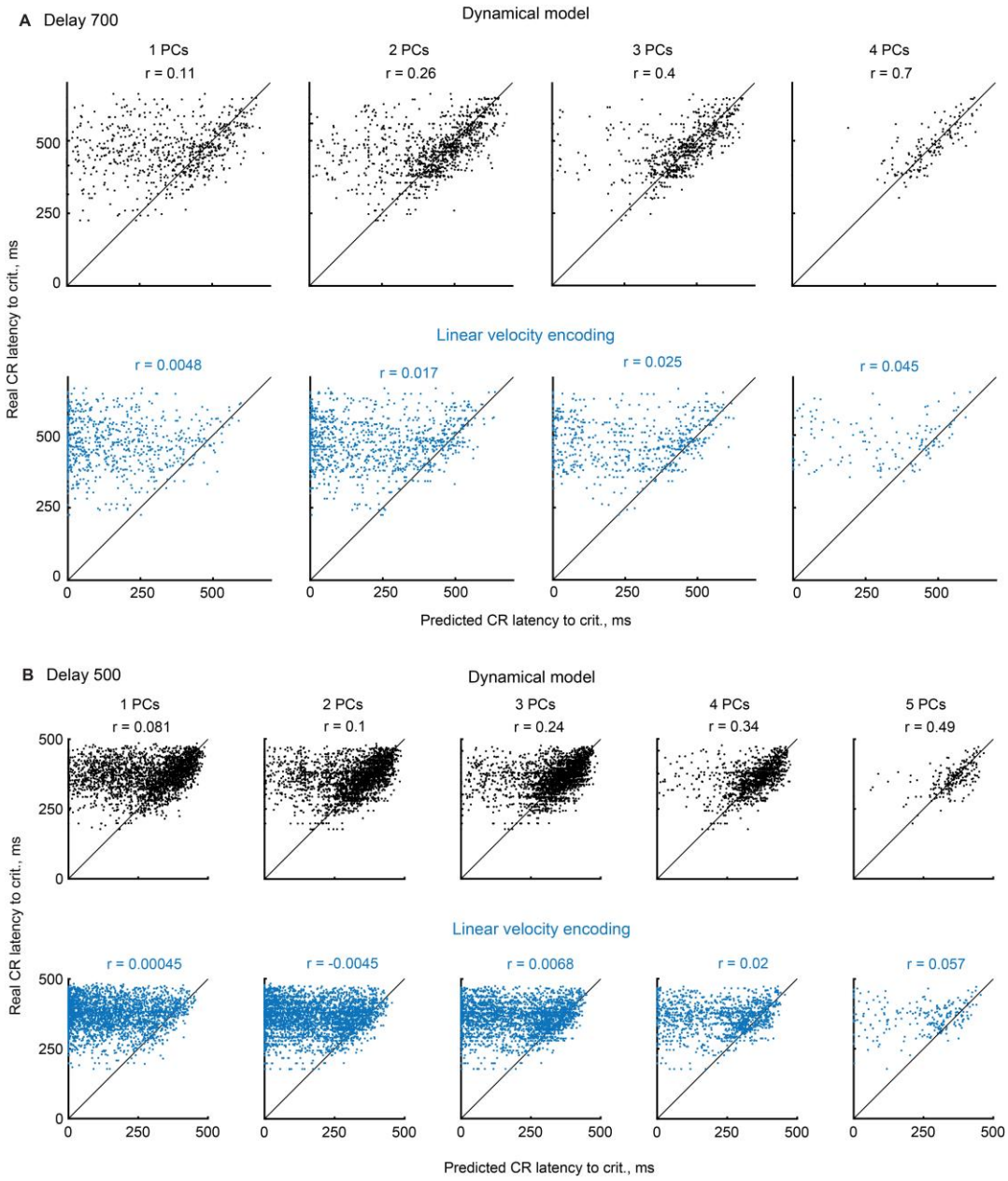

Fig. 2.11. Single trial prediction of CR latencies.
A) Performance of a dynamical model (upper row, black color) and linear velocity encoding model (bottom row, blue color) for Delay 700 protocol. Panels are arranged in columns by the number of eyelid PCs used for prediction. Each dot represents a single trial with real versus predicted CR latencies to 1 mm criterion plotted against each other. Corresponding concordance correlation coefficient is indicated for each panel. Diagonal is shown by a black line. B) The same for Delay 500 data.

2.5 Discussion

For the present analysis I used only eyelid PCs, classified based on the presence of complex spikes driven by the US. This selection criterion is likely to be crucial for the precise relationship between PCs simple spike trains and behavior I observed. One of the advantages of the eyelid conditioning paradigm is the direct mapping of CS and US inputs onto the cerebellar circuitry, with US activating climbing fibers (though note of studies reporting CS onset to drive PCs complex spikes in mice (Ohmae and Medina, 2015)). Therefore, by restricting my area of interest only to PCs with US driven complex spikes, I am able to isolate the population of output neurons that received a signal driving cerebellar learning specifically during eyelid conditioning paradigm. Consistent with this logic, our previous analysis showed that the activity of non-eyelid PCs is related weakly to such CR features as onset time and amplitude and at best it able to differentiate between CR and nonCR trials (Halverson et al., 2015). Studies involving different cerebellar behavioral paradigms also highlight the importance of PCs complex spike tuning for subsequent simple spikes analysis. For example (Herzfeld et al., 2015) analyzed simple spikes of PCs recorded in the oculomotor vermis region of the cerebellum in rhesus monkeys performing saccades. The authors found that only when PCs simple spikes were grouped according to the complex spikes tuning to saccade direction, the overall PCs population activity matched the behavioral responses. If the method of analysis I presented in this chapter will be used in future to study the

relationship between PCs simple spikes and motor responses in other behavioral paradigms, the importance of complex spikes tuning in individual PCs should not be overlooked.

The results of the analysis above demonstrate a way to predict the CR time-profile from PCs spike trains on a single trial. There are several possible direct applications of these results. First, this analysis could be applied to experimental setups where recordings of PCs activity are done without behavioral output due to an animal being immobilize (Jirenhed et al., 2007; Svensson et al., 2010; Jirenhed and Hesslow, 2011). Second, it can be directly applied to cerebellar simulations, ranging from large-scale spiking networks containing over 10^6 neurons ((Li et al., 2013) and section 3.4.5 of Chapter 3) to simple firing rate models containing a single PC. By using this analysis the predicted behavioral readout from the simulation can be compared to experimental data, enhancing the predicted power and verifiability of conclusions drawn from the simulation. Application to other cerebellar learning paradigms should be mindful of possible cross-species differences resulting in non-identical coefficients in the dynamical model equation (2.7) describing the properties of the eyelid plant. Thus, if behavioral data comes from a different species other than rabbits, ideally the analysis described above should be repeated on that data before a direct comparison to simulation results can be made.

I have shown here that the relationship between PCs spike trains and behavioral responses is best described via a dynamical model where PCs do not explicitly encode

any single kinematic variable of a CR. I would like to note however that the assumption of linear velocity encoding also produced satisfactory results, consistent with published studies (Medina and Lisberger, 2007, 2009; Heiney et al., 2014). Therefore, I would argue that my results do not dismiss a common notion that PCs linearly encode velocity of the movement but rather highlight that such a notion is likely a simplification of the true underlying relationship. A similar idea has been recently proposed in the motor cortex literature (A. Russo, B. London, S. Perkins, 2016), where authors used a complex motor task with rhesus monkeys navigating a virtual environment using a hand to pedal either forward or backward. The study showed that while the relationship between neuronal activity and kinematics was different for two movement directions, EMG activity paralleled neural data and could be successfully decoded from it. These result support the notion that at least during the execution of a movement motor areas do not need to encode specific aspects of the movement like velocity, but simply provide a signal to downstream targets to successfully execute the movement.

CHAPTER 3

Cerebellum implements probabilistic binary choice as adaptation to uncertain inputs

3.1 Abstract

Noise and variability are inherent and unavoidable features of neural processing. Despite that, brain systems function well, suggesting an existence of adaptations. Here I report a novel adaptation that the cerebellum implements to maintain a correct and adaptive response in the face of ambiguous inputs. I found that under these conditions the cerebellum employed a probabilistic binary choice: the probability of a behavioral response gradually changed with the similarity between current and trained inputs, but the size of the response remained constant. That way the cerebellum kept responses adaptive to a trained input corrupted by noise, while minimizing false responses to novel stimuli. Recordings and analysis of PC activity showed that the binary choice is made in the cerebellar cortex. Results from large-scale simulations suggest that internal feedback from the cerebellar nucleus back to cerebellar cortex plays a critical role in the implementation of binary choice.

3.2 introduction

Noise and variability are inherent and unavoidable features of neural processing. Despite that, brain systems function well, suggesting an existence of adaptations. A number of studies have explored adaptations that sensory systems use to efficiently decode the stimulus from neural activity (Field, 1987; Tkacik et al., 2010; Zylberberg et al., 2016). Motor systems, however, are presented with a different challenge – what should be a correct output given an ambiguous noisy input? Imagine a following scenario: a person had learned to perform a motion, e.g. catch a ball, by extending an arm by the right amount. When faced with an input only partially similar to the incoming ball, what should be a proper motor response? To decrease the amount of arm extension will automatically result in missing the ball and a movement error. However always performing a learned movement is also not adaptive, since partial similarity in the example above could mean a novel input. Indeed, in this chapter I show that under these conditions the cerebellum implements a binary choice*: the cerebellum scales the probability of producing a response with the similarity between trained and presented inputs; however if the response is made – the amplitude maintains an adaptive,

* The word "binary" is used here more in illustrative way than in an exact sense. There is a natural spread of CR amplitudes to trained input even in well-trained animals. Therefore by "binary" we mean that the distribution of CR amplitudes should be the same in response to trained and probe inputs, rather than every CR is expected to be exactly target-sized.

previously-trained value. This way the cerebellum maintains a correct and adaptive motor response even in the face of ambiguous inputs.

3.3 Methods

3.3.1 Initial training.

During initial training, subjects were given daily eyelid conditioning sessions comprised of 12 blocks of 9 trials each. For subjects trained to produce full-sized CRs, each block consisted of 1 conditioned stimulus-alone (CS-alone) trial and 8 paired trials (CS+US). For subjects trained to make half-sized (3 mm) CRs, a block consisted of 1 CS-alone trial and 8 trials which were either a paired or CS-alone trial, depending on the CR amplitude before the US delivery. Following published training procedure (Kreider and Mauk, 2010), eyelid position was monitored throughout the presentation of the CS, allowing to calculate the CR amplitude before US presentation. If CR amplitude was equal or larger than the target (3 mm) 10 ms before US presentation, the US was skipped at that trial. Otherwise, if the CR size was too small, the US was presented at the end of the CS. All subjects were trained at an inter-stimulus interval (ISI) of 500 ms. Each conditioning chamber was equipped with a speaker that was connected to a stereo equalizer and receiver which were connected to a computer that generated the tone. For subjects trained using tone as the CS, the CS was set as a 1 kHz, 500ms, 75 dB sinusoidal tone with a rise and fall time of 5 ms to avoid audible clicks from the speaker. For

subjects trained with mossy fiber stimulation, the trained CS was a constant frequency pulse train of cathodal current pulses (100 Hz, 500 ms, 0.1 ms pulse width, 100-150 μA). The period of initial training was not fixed, but rather continued for each rabbit until both CR percentage was high (CR% > 90%) and CR amplitudes were robust and near the target amplitude. Typically, such pretraining lasted for 10 sessions from a naïve state.

3.3.2 Probe sessions.

After the initial training was complete, the subjects were switched to probe sessions. Each probe session included 30 CS-alone probe trials interspersed with 80 paired trials. For subjects trained to half-sized CRs, only a portion of the 80 trials were paired depending on the CR amplitude on a given trial as described above. The purpose of the paired trials during the probe sessions was to maintain the high level of CRs to the trained stimulus. It would be ideal to include as many CS-alone probe trials in the session as possible, however, it is important that probes were infrequent enough that they did not change responding during the session due to extinction. Our studies showed that for well-trained subjects having 30 probe trials per session fulfills this requirement.

I implemented three probe protocols that systematically altered the probe stimulus from the trained CS in different ways: 1) *Short Probes:* Probe durations ranged from 50 ms to 450 ms in 50 ms increments. 2) *Frequency Probes:* probe duration was not changed from the trained CS, but stimulation frequency was decreased from the original 100 Hz down to 50 Hz in 10 Hz increments. 3) *Competing Stimulus Probes:* The

temporal pattern of the trained CS (duration and frequency) was unaltered. However the current intensity delivered through stimulation electrodes was adjusted to gradually shift the population of activated mossy fibers, maintaining only a partial overlap with mossy fibers used for training (Fig. 3.1 E). Specifically, the current intensity delivered through the electrode used in training was decreased on probe trials, but an additional 'competing' input through the second electrode, which was not used during acquisition, was delivered to maintain approximately the same total amount of activated mossy fibers by the trained CS and the competing stimulus probes. We used here a simple assumption that the number of activated mossy fibers is determined by the area of current spread, the diameter of which is proportional to the current intensity.

$$N \sim d^2 \sim I^2$$

Where N is the number of activated mossy fibers and I is current intensity applied through the electrode. Such an assumption was our best approximation to maintain the same number of activated mossy fibers across all probe types. Only one type of probe protocol was used in a single probe session. However, the same subjects were often used in sessions with different probe protocols. Typically, 5 probe types (e.g. 50 Hz, 60 Hz, …, 90 Hz) were presented during the session, with 6 probe trials per probe type per session. Probe sessions were repeated several times for each rabbit in order to collect a sufficient amount of probe data. Some probe sessions also included CS-alone trials with the trained CS. Responses to the trained CS were extracted from these sessions and from

CS-alone trials during normal training sessions that directly preceded, followed or were in-between the sequence of probe sessions. The period of initial training was not fixed, but rather continued for each rabbit until both CR percentage was high (CR% > 90%) and CR amplitudes were robust and near the target amplitude. Typically, pre-training lasted for 10 sessions from a naïve state.

3.3.3. In-vivo recordings.

Only well-isolated single units were used for analysis. In the previous chapter I demonstrated that the activity of eyelid PCs is sufficient to predict CRs during conventional eyelid conditioning sessions. If a computation underlying the behavioral phenomenon described below occurs in the cerebellar cortex, I should observe the same phenomenon in eyelid PCs activity. Therefore eyelid PCs were the main target of in-vivo tetrode recordings performed during probe protocols sessions. Only well-isolated single unit were used for analysis. I recorded 491 single units during the binary choice sessions. Out of those, 82 units were classified as eyelid PCs and 116 units as non-eyelid PCs. All single unit analysis reported in this chapter was performed on simple spikes from eyelid PCs. Recordings where the unit was lost at any time during the session were not included in the analysis.

3.3.4 Binary index measure.

For every probe protocol I calculated the dependence of CR amplitude on CR probability, where each probe type contributed a single point on the plot (Fig. 3.2 C, Fig. 3.8 A, B). In the ideal implementation of binary choice, CR amplitude would be constant over the whole range of CR probabilities, i.e. CR amplitude would be independent from CR probability (Fig. 3.9 B, black squares). In order to quantify the degree of dependence of CR amplitude on CR probability I calculated the binarity index (BI). BI was defined as the area between the CR amplitude curve (Fig. 3.9 A right panel, red line) and the diagonal line (grey), connecting a point corresponding to the trained CS with the point of origin. Red shaded area in Fig. 3.9 A illustrates this procedure. The measure was normalized so that the ideal BI (black shaded area) is equal to one. In order to evaluate the expected mean and standard deviation of BI, 1000 samples of 50 random trials per each probe type within a protocol were drawn and for each sample BI was calculated as described above. For each sample, if the fraction of CRs was less than 10% for a particular probe, it was not included in the BI calculation.

The BI measure was also used to evaluate the independence of PCs responses from CR probability. The BI definition there was identical, except CR amplitude was replaced with 'relative PC response' (Fig. 3.8 D, E). Relative PC response was defined as the ratio between 1) mean decrease from baseline in PC spike count on CR trials in

response to trained CS and 2) decrease in spike count on CR trials in response to probe stimuli.

3.3.5 ROC analysis.

Because the data I observed resembled probabilistic decision making, I employed analyses from signal detection theory which are commonly used in decision making studies. To quantify the relationship between eyelid PCs activity and behavioral responses, I employed a receiver operating characteristic (ROC) analysis. For each probe stimulus type, the trials were grouped into two groups based on the behavioral responses: CR trials or nonCR trials. Note that such grouping only assures the presence of a CR on a CR trial, but does not assure specific CR amplitude. By the design of the session there were only 6 trials per probe type per session (and therefore per individual PC), I did not compute an ROC curve separately for each cell. Instead, I combined spike counts from all recorded eyelid PCs on different sessions into a single "grand" distribution for each probe. A session was added into the grand distribution only if at least one CR and one nonCR was present on a given probe type. For each recorded eyelid PC, spike counts were normalized to the average spike count during the pre CS baseline. Normalization allowed us to reduce the influence of the baseline firing variance in different PCs on the shape of the "grand" spike count distribution. ROC curves were calculated from the spike counts of the grand distributions, an area under the curve (AUC) was calculated to measure the probability of correctly predicting the behavioral outcome (CR versus

nonCR) from PCs spike count on a random trial. However, the resultant AUC from the grand distribution can only be less than the mean across AUC values computed for each cell separately. Therefore, AUC values reported here, referred to as choice probabilities, represent a lower estimate of the mean choice probability across the eyelid PCs population. To measure an overall ability to predict the behavioral outcome, PCs spike counts were calculated over 700 ms from the CS onset window. To quantify a finer temporal structure of choice probability, I calculated spike counts over non-overlapping 100 ms time windows and used an ROC analysis on each window separately in the same fashion as described above. In order to align spike trains on nonCRs trials to CR onset time, for each nonCR trial I randomly draw 1000 times with replacement from the CR onsets distribution, calculated an aligned to CR onset spike train for each draw and finally computed the mean aligned to the CR onset spike train. Error-bars shown for AUC values represent 95% confidence intervals calculated by bootstrapping 1000 samples. A permutation test was used to measure significance above chance (5000 samples).

3.3.6 Computer simulations

Implementation of large-scale computer simulation experiments for this project was performed by Evan DeLord. Design of experiments and data analysis was performed by me. We used a large-scale simulation of the cerebellum (over 10^6 simulated neurons), composed of conductance-based, single compartment, spiking representations of neurons.

The properties of the simulation were intended to emulate the synaptic organization and physiology of the cerebellum (Eccles, 1967). Connectivity between neuronal types in the simulation mimicked known cerebellar circuitry, for details see (Li et al., 2013). As an approximation to the ratio of cell types within the cerebellum, the simulation was comprised of: 1024 mossy fibers, 1048576 (2^{20}) granule cells, 1024 Golgi cells, 128 basket cells, 512 stellate cells, 32 PCs, 8 deep cerebellar nuclei cells and 32 climbing fibers. These neurons were interconnected to emulate a parasagittal stripe, where all PCs receive input from climbing fibers of the same type – that is, all PCs in the simulation were "eyelid" PCs per the definition used for in-vivo recordings.

A new addition in the current version of simulation is the presence of deep cerebellar nucleus axon collaterals (DCNcol) that projected as mossy fibers back to the granule-Golgi cell network. We implemented 30 DCNcol in the simulation. DCNcol had the same connectivity to Golgi and granule cells as normal mossy fibers, but the spiking activity conveyed by each of them was a copy of one of the eight DCN neurons (chosen randomly at simulation initialization). In this way we have incorporated axon collaterals only from glutamatergic DCN neurons(Houck and Person, 2015) , leaving contributions from GABA/glycinergic neurons (Ankri et al., 2015) aside.

The state of the network was updated every millisecond. The only external inputs to the simulation were the spiking activity of 1024 mossy fibers and 32 climbing fibers, one per each PC. Synaptic plasticity sites were implemented at the granule cell-to-PC

synapses and at the mossy fiber-to-DCN synapses. A granule cell-to-PC synapse underwent LTD or LTP every time a granule cell fired a threshold burst of spikes: LTD occurred if this burst fell within a window between 200 ms and 100 ms prior to a climbing fiber input to the PC, otherwise LTP occurred. Mossy fiber-to-DCN synapses active within a time window of an abrupt pause in PCs activity underwent LTP, whereas those active during a strong increase of PCs activity underwent LTD.

To train the simulation, 25 mossy fibers were assigned to be CS driven. These mossy fibers changed their firing rate from 5 Hz to 60 Hz for the duration of the CS. Spike times for each mossy fiber were determined from a Poisson distribution with a given mean. The only exception was the CS driven mossy fibers in simulations trained for the frequency probes experiment. To directly replicate these experiments in the simulation, spike times during the CS were set to a constant 100 Hz frequency. To keep the strength of the CS input comparable with other probe protocols, the number CS driven mossy fibers was reduced in these experiments by half. An excitatory conductance was applied to the climbing fibers to represent the presentation of the US. Both types of simulations: with and without DCNcol, were trained for 2000 paired trials at 500 ms ISI with a 5 s inter-trial interval. In both types of simulations the performance reached a plateau after 400-500 trials. After the training period the plasticity was frozen so that synaptic weights would not change during subsequent CS-alone probe trials. We tested computer simulations with the same probe protocols as used in experiments. Probe trials were delivered with a 5 s inter-trial interval. Each probe type was presented for

over 200 trials with simultaneous recordings of simulation PCs activity and eyelid behavioral data, or over 2000 trials if only eyelid behavioral data was collected. The simulated eyelid position was calculated as a smoothened sum of all DNC neuron action potentials with a decaying time constant that was passed through a threshold function. The amplitude of the simulated eyelid position was scaled such that the maximum CR amplitude over the 2000 trials training period corresponded to a 6 mm eyelid closure. All simulations were run on a custom-built computer housing 8 graphics cards (NVidia, interfaced through CUDA) totaling 12,286 GPU cores.

Analysis of simulation data was implemented in a way that paralleled analysis of experimental data. For analysis of simulation PCs activity, I randomly sampled different PCs over 6 trials of each probe type per cell to replicate the session design used in experiments.

3.4 Results

3.4.1 Behavioral experiments design

Because these studies require strong control over inputs to the cerebellum I employed eyelid conditioning as a cerebellum-dependent behavior that permits such control (Steinmetz et al., 1989; Kalmbach et al., 2011). In experiments described below, inputs to the cerebellum were controlled via electrical stimulation of mossy fibers

(Steinmetz et al., 1985; Kalmbach et al., 2011). The stimulation was delivered through electrodes implanted in the middle cerebellar peduncle (Fig. 3.1 A), comprised solely of mossy fiber axons projecting to the cerebellum. Subjects were trained with a 500 ms long 100 Hz pulse train as the conditioned stimulus (CS). I will refer to this stimulus as the trained CS through this section. For most experiments, subjects were trained to produce full sized CRs close to 6 mm, but a subset were trained to elicit half-sized (3 mm) CRs (Fig 3.1 B). The training continued until animals showed close to target CR amplitude on nearly every trial. Each subject then received numerous sessions in which a subset of the trials involved unreinforced probe inputs that were, to varying degrees, different from the trained input. I used three protocols to alter probe stimulus from the trained CS in three different ways: frequency probes, short probes and competing stimulus probes (Fig. 3.1 C-E), where I systemically altered probe stimulus temporal structure, duration or overlap with the trained CS respectively.

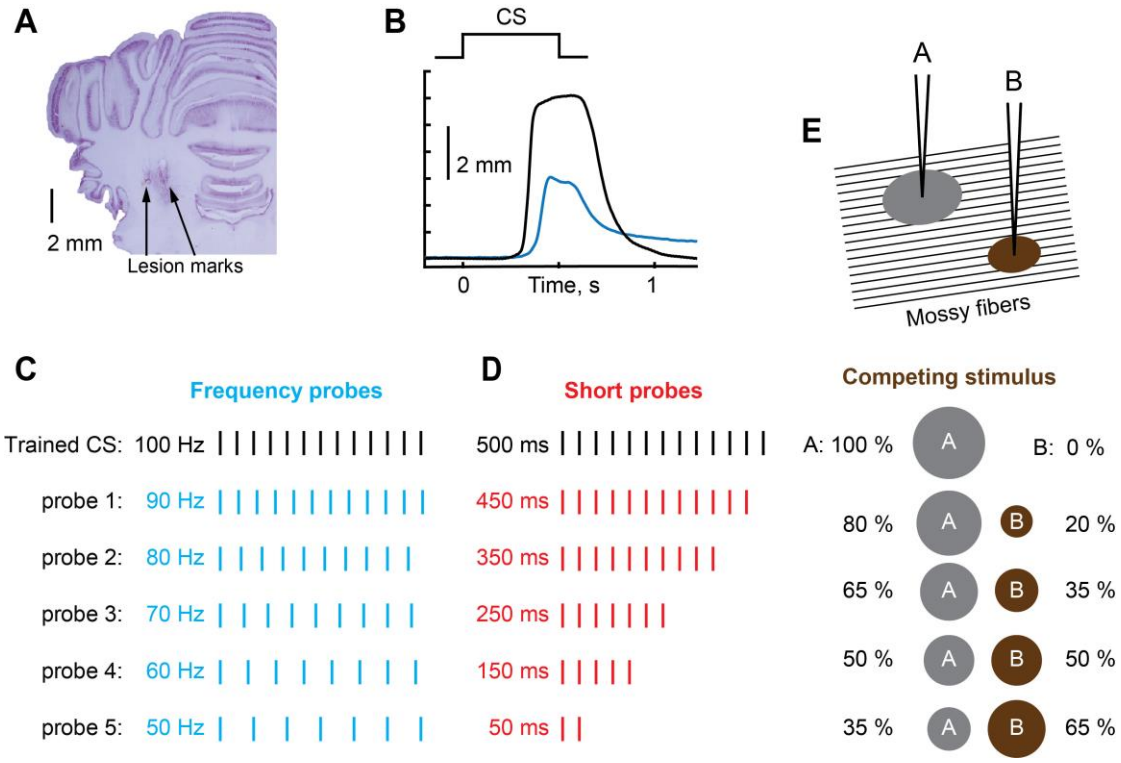

Fig. 3.1. Schematics of eyelid conditioning training and probe protocols.
A) Coronal cross-section of histology with lesion marks from stimulating electrodes in middle cerebellar peduncle B) Example eyelid response profiles on CS-alone trial (no US) from subject trained to produce either a full sized eyelid closure (black line) or half-sized (blue line). C-E) Schematics of probe inputs on three different probe protocols. In each case CS used for training (500 ms 100 Hz pulse train) is shown in black. C) Frequency probes. The length of stimulus is kept at 500 ms, but frequency is systematically decreased. D) Short probes. Frequency is kept at 100 Hz, but only portion of stimulus length is presented. E) Competing stimulus. Two separate stimulating electrodes were implanted into the middle cerebellar peduncle spaced 1 mm laterally. Only electrode A was used for training. During probe trials neither frequency or length of the stimulus were changed, but rather the current applied on electrode A was decreased and correspondingly current on electrode B was increased from zero to keep the number of total activated mossy fibers approximately constant. Such manipulation should result in a gradual shift of the overlap between mossy fibers activated by probe and trained CS. The area of current spread is illustrated as a grey circle for electrode A (used to deliver trained CS) and a brown circle for electrode B (competing stimulus).

3.4.2 Responses to probe stimuli reveal probabilistic binary choice

The essential finding, observed in all three probe protocols, is that although the probability of a CR decreased as the probes were made more different than the trained CS, the amplitude of the CRs remained at the previously trained level. This is evident in the example behavioral responses from a single subject, elicited by either the trained CS or 70 Hz (frequency) probes shown in Figure 3.2 A. The trained CS elicited an eyelid CR on 92% of the trials and the distribution of CR amplitudes (in black) showed a clear peak at 6 mm, the full eyelid closure targeted by the previous training. In turn, 70 Hz probes elicited CRs in fewer trials (46%) and the distribution of CR amplitudes (in cyan) was bimodal – with one mode corresponding to non-responses and the second mode to the full-sized target responses, as with the trained CS. The summary plots across all animals, probe protocols and probe types are shown is Fig. 3.2 B and C, where I systematically varied stimulus parameters in each protocol. Here lines in cyan, brown and red correspond to frequency probes, competing stimulus and short probes respectively, black dots represent responses to the trained CS. For all three protocols, CR probability gradually decreased as the probe stimulus deviated from the trained CS in any direction of parameter space (Fig.3.2 B, linear fit $R^2 > 0.72$ for all protocols). The gradual change in CR probability was not the result of averaging across animals and was apparent within every animal. However, CR amplitudes were all-or-none in all three protocols and showed virtually no change with different probe types or CR probability (Fig. 3.2 C, linear fit $R^2 < 0.02$ for all protocols). In other words, even for probes that

elicited CRs with less than 50% chance, if CRs happened, - their amplitude was full-sized, as with the trained CS.

Eyelid CR amplitudes could be all-or-none for the trivial reason that eyelid responses are inherently all-or-none. Although previous studies (Kreider and Mauk, 2010) suggest this is not the case, I additionally tested this by training a group of animals to produce CR amplitudes with a target size of 3 mm, corresponding to half of a full eyelid CR. When these animals were tested with frequency probes (blue lines in Fig. 3.2 B, C), CR probability decreased as the probes became more different from the trained input (linear fit $R^2 = 0.78$), but CR amplitudes remained relatively constant to the trained 3 mm size, independent of CR probability ($R^2 = 0.03$). These data demonstrate that binary choice selects between a non-response and a response of the previously-trained amplitude, and is not simply a consequence of an inherently all-or-none response system.

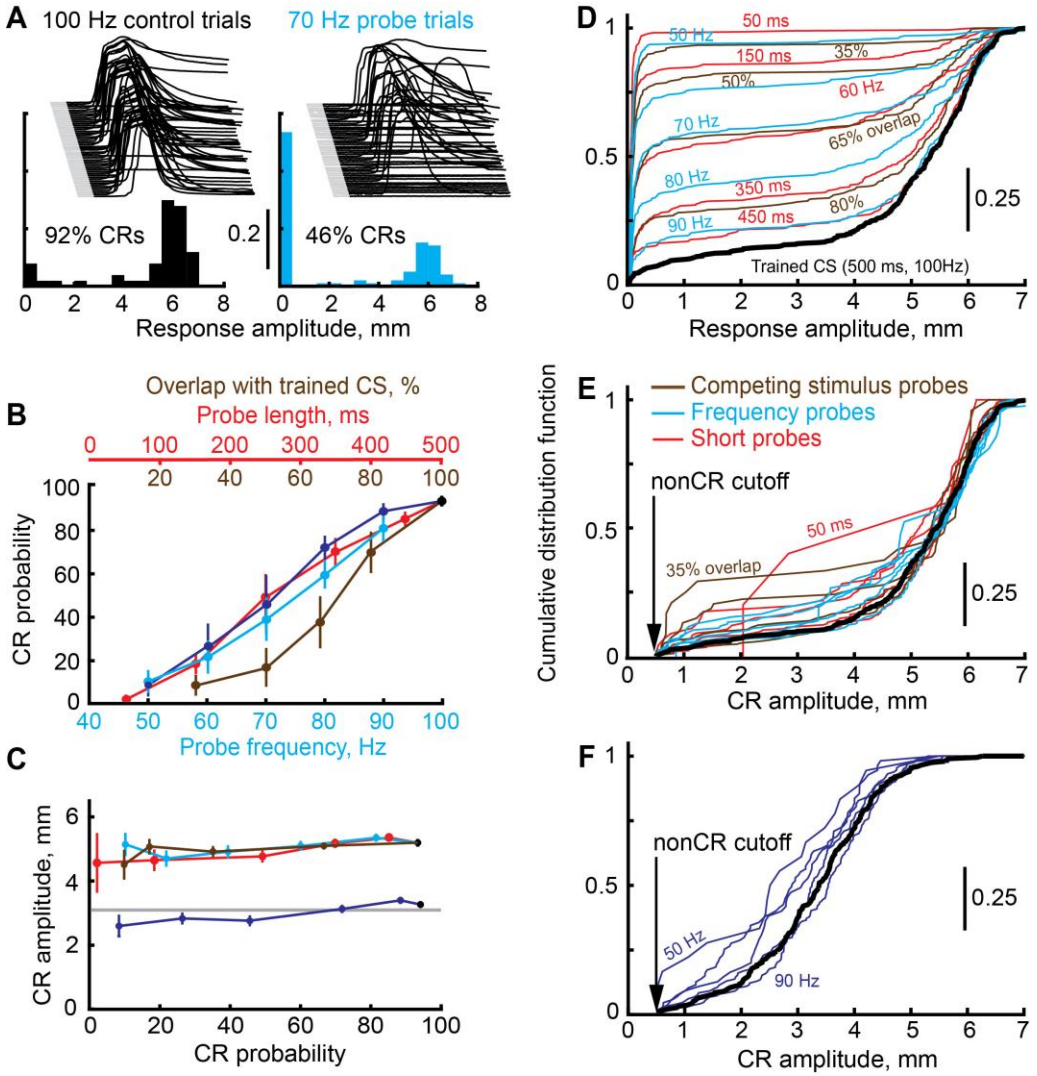

Fig. 3.2. Behavioral summary of binary choice.
A) Example eyelid responses and frequency distribution of response amplitudes to trained CS (100 Hz, black, left), and to 70 Hz frequency probes (cyan, right). In about half of trials 70 Hz probe resulted in non-CR, but CR amplitudes were all-or-none when they happen. B) CR probability decreases as probe stimuli become more different from the trained input. Data is shown for all three protocols: competing stimulus probes (brown), short probes (red) and frequency probes (cyan for animals trained to 6 mm CRs, blue – to 3 mm CRs). C) Mean CR amplitude as a function of CR probability remained constant in all three probe protocols. D) CDFs of distributions of response amplitudes to each probe (color-coded) or trained input (black). E) Same as D, but with non-CRs removed from each distribution. F) Same as E for rabbits trained to 3 mm target CR amplitude.

There is a degree of natural variability in CR amplitudes, even for the responses to the training stimulus. If binary choice operates to maintain CR amplitude at the previously trained size, the distributions of CRs amplitudes to either probe stimuli or trained CS should be statistically indistinguishable after nonCRs are omitted. CDFs of all eyelid response amplitudes (CRs and nonCRs) for every probe protocol and type are shown in Fig. 3.2 D. In each case, the differences between responses to the trained CS (black line) and probes were highly significant, except for probes with parameters closest to the trained CS (two-tailed Kolmogorov-Smirnov test, Table 3.1 for p values). After omitting nonCR trials from the distributions, all CDFs collapsed onto the trained CS CDF (black line), as shown in Panel E, and were not statistically different from it ($p>0.1$ for 12 probes, $p>0.05$ for 2 probes, without a correction for multiple comparisons). The same analysis done on animals trained to half-sized, 3 mm CRs yielded the same conclusions (Fig. 3.2 F).

Frequency probes

	50 Hz probe	60 Hz probe	70 Hz probe	80 Hz probe	90 Hz probe
All responses	$p = 3 \cdot 10^{-128}$	$p = 1 \cdot 10^{-68}$	$p = 1 \cdot 10^{-37}$	$p = 4 \cdot 10^{-13}$	p = 0.05
Without nonCRs	p = 0.32	p = 0.07	p = 0.14	p = 0.17	p = 0.34

Short probes

	50 ms probe	150 ms probe	250 ms probe	350 ms probe	450 ms probe
All responses	$p = 1 \cdot 10^{-168}$	$p = 3 \cdot 10^{-87}$	$p = 5 \cdot 10^{-30}$	$p = 3 \cdot 10^{-7}$	p = 0.12
Without nonCRs	p = 0.65	p = 0.26	p = 0.051	p = 0.99	p = 0.42

Competing stimulus

	35% overlap	50% overlap	65% overlap	80% overlap
All responses	$p = 1 \cdot 10^{-101}$	$p = 3 \cdot 10^{-77}$	$p = 1 \cdot 10^{-33}$	$p = 7 \cdot 10^{-6}$
Without nonCRs	p = 0.11	p = 0.20	p = 0.39	p = 0.35

Frequency probes, 3 mm target CR amplitude

	50 Hz probe	60 Hz probe	70 Hz probe	80 Hz probe	90 Hz probe
All responses	$p = 7 \cdot 10^{-66}$	$p = 1 \cdot 10^{-38}$	$p = 3 \cdot 10^{-22}$	$p = 5 \cdot 10^{-5}$	p = 0.46
Without nonCRs	p = 0.07	p = 0.21	p = 0.03	p = 0.47	p = 0.53

Table 3.1. Two-sample Kolmogorov-Smirnov test, comparison between CR amplitude distributions to probe and trained inputs.

3.4.3 Purkinje cells activity also demonstrates binary choice

The use of direct electrical stimulation of mossy fibers as the training and probe stimuli eliminates potential contributions to binary choice behavior from processes upstream of the cerebellum. If binary choice is computed within the cerebellum, then the responses of cerebellar output neurons should also reflect the phenomenon. I used *in vivo*

recordings to determine whether the responses of PCs – the principle neurons and sole output of the cerebellar cortex – show binary choice. Six subjects were prepared with chronically implanted tetrode microdrives in the region of cerebellar cortex previously shown to be necessary for the expression of eyelid CRs (Fig. 3.3 A) (Perrett and Mauk, 1995; Garcia et al., 1999). I targeted, and restricted analysis to "eyelid" PCs, identified by the presence of US-evoked complex spikes (Halverson et al., 2015; Ohmae and Medina, 2015) (Fig. 3.3 C, D, Fig. 3.4 A). That way, the analysis was done only on output neurons that received a learning-related signal specific to our behavioral paradigm. During recordings sessions I implemented frequency and short probes protocols with a mossy fiber stimulation CS as well as short probes with a 1 kHz tone CS. Tone sessions were done to test if I will observe the binary choice phenomenon with natural stimuli. Fig. 3.4 B shows an example raster plot from an eyelid PC along with the behavioral responses on the left. Green dots indicate CR onset times on different trials. All trials here are CS-alone frequency probes trials.

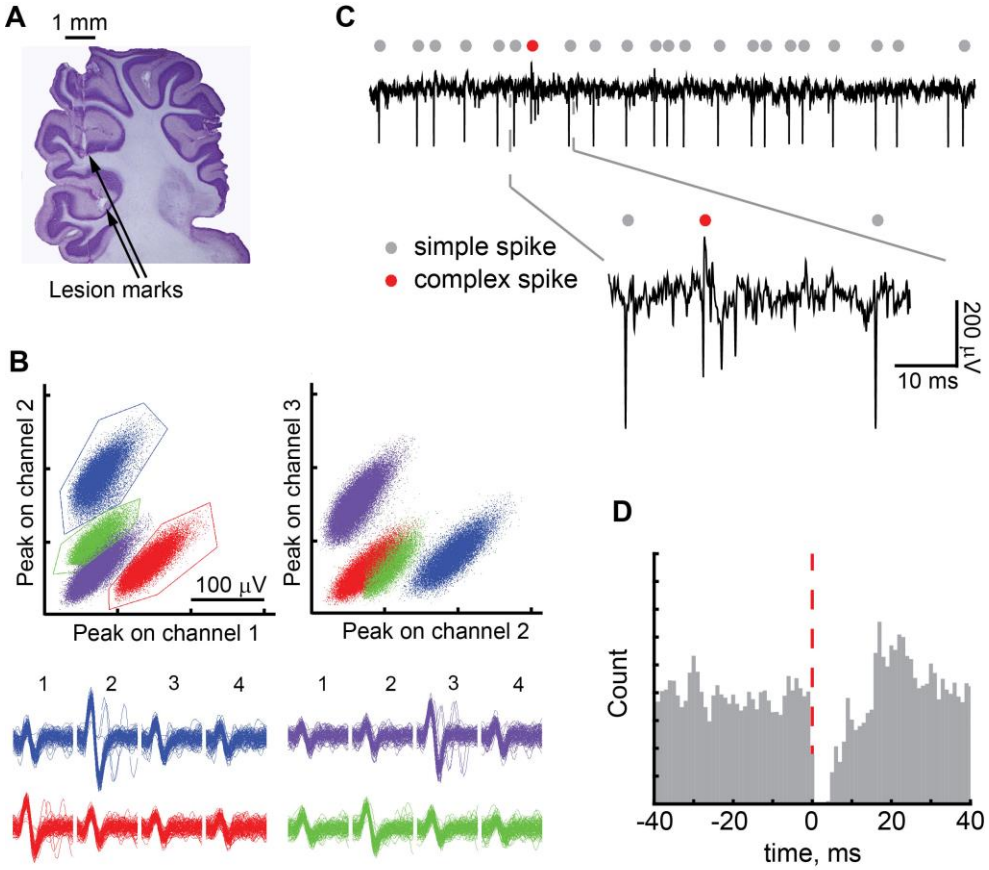

Fig. 3.3. Isolation of single units and eyelid Purkinje cells from tetrode recordings.
A) Sagittal view of histology with tetrode tracks in the cerebellar cortex B) Two top panels show four isolated single units in two cluster-cutting projections (peak on tetrode's channel 2 versus channel 1 and peak on channel 3 versus channel 2). A hundred overlaid waveforms from each unit (color-coded) recorded on each channel is shown at the bottom. C) Example continuous recording from a tetrode's channel with a highly isolated eyelid PC. Grey dots indicate times of simple spikes, red dots indicated complex spikes. Times of simple and complex spikes were found from cluster-cutting procedure. A zoomed in portion with simple and complex spike waveform is shown on the right. D) Spike-triggered average of simple spikes on complex spikes, demonstrating a post complex spike pause..

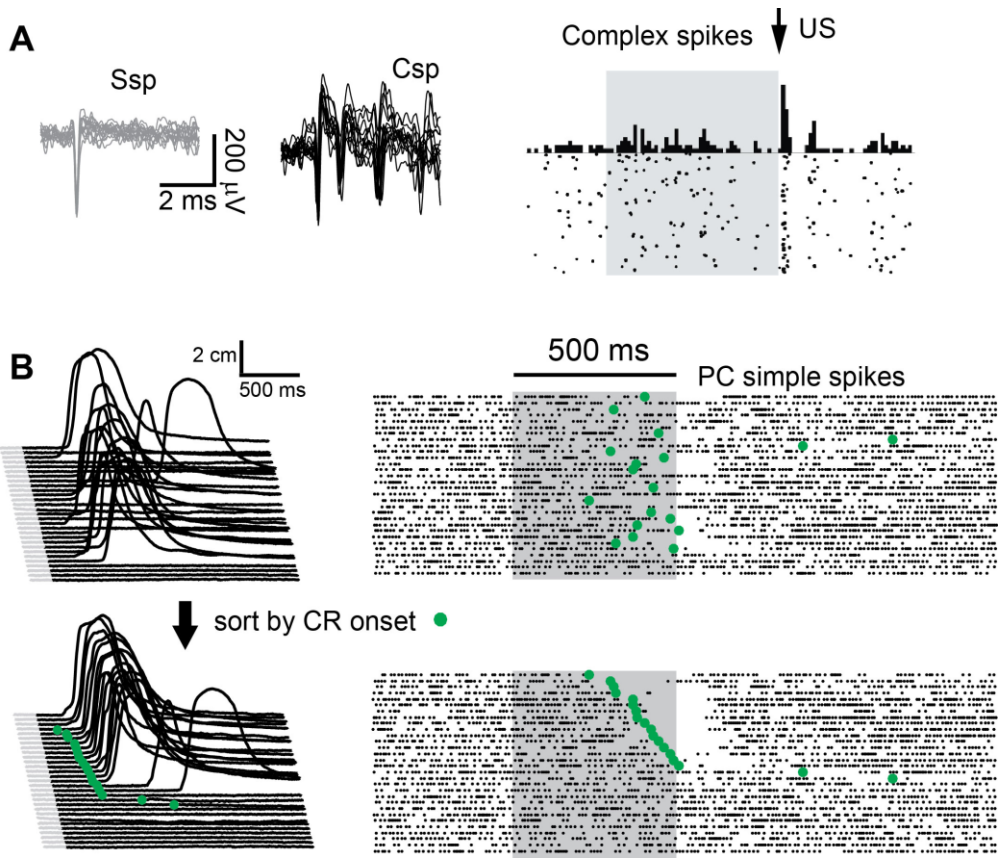

Fig. 3.4. In-vivo recordings from Purkinje cells during binary choice sessions.
A) Twenty overlaid waveforms of simple (grey) and complex (black) spikes are shown on the left, PSTH of US-evoked complex spikes from eyelid PC is shown on the right. B) Behavior and raster plot of PC simple spikes during CS-alone frequency probes trials. Trials are sorted either in order of occurrence (top) or according to CR onset time (green dots, bottom). CS duration is indicated by the grey shaded area.

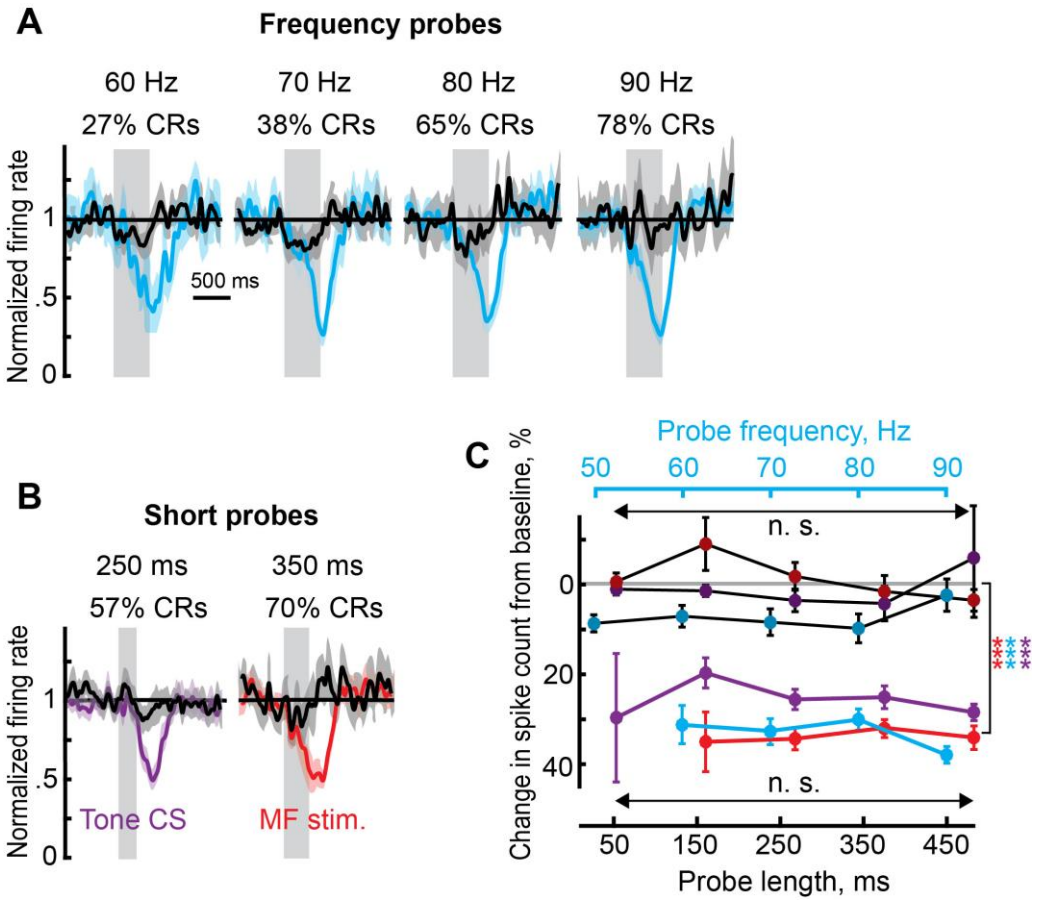

Fig. 3.5. Eyelid Purkinje cells responses during binary choice sessions.
A) Average firing rate of eyelid PCs on CR (cyan) or non-CR trials (black) during frequency probes. Grey region represents probe duration. Shaded regions indicate 95% confidence intervals. B) Same as D, but individual examples for short probes, using either mossy fiber stimulation (red) or 1 kHz tone (violet). C) Average spike count of eyelid PC activity on CR (color-coded by probe protocol) and non-CR trials (black lines, darker colored points to indicate the protocol).

Sorting trials based on CR onset times, shown at the bottom panel, reveals a tight relationship between eyelid PC activity and behavior, even on a single cell – single trial level. On nonCR trials shown at the bottom of the raster plot, eyelid PC activity barely deviated from the baseline level. On trials with CRs present, the start of decrease in eyelid PC firing rate tightly matched CR onset time on that trial. Average eyelid PC firing rate profiles on CR trials (cyan for frequency probes, red and violet for short probes with mossy fiber stimulation or 1 kHz tone as a CS, respectively) and nonCR trials (black) are shown in Fig. 3.5 A, B. Responses of all recorded eyelid PCs were combined across trials with the same probe type. In all cases, eyelid PCs firing rate showed a clear difference on CR versus nonCR trials, independently of CR probability or probe protocol. To quantify that effect, I calculated average spike counts on CR and nonCR trials over a 700 ms window from CS onset, shown in Fig. 3.5 C. Indeed, for all protocols there was highly significant difference between spike counts on CR versus nonCR trials (two-way ANOVA, $p < 10^{-7}$ for all protocols, see Table 3.2). Eyelid PCs responses on CR or nonCR trials did not change with probe type or CR probability (two-way ANOVA, $p > 0.2$ for all protocols), paralleling behavioral binary choice. Occasional differences in the amount of PCs decreases across protocols are likely due to sampling from different eyelid PCs, where the amount of CR related decrease can vary from cell to cell.

	Probe type	CR/nonCR	Interaction
Frequency probes	$F_{3,414} = 0.08$ $p = 0.97$	$F_{1,414} = 152.1$ $p = 6 \cdot 10^{-30}$	$F_{3,414} = 2.25$ $p = 0.08$
Short probes, MF. stim. CS	$F_{3,468} = 0.53$ $p = 0.66$	$F_{1,468} = 58.4$ $p = 1 \cdot 10^{-13}$	$F_{3,468} = 0.66$ $p = 0.58$
Short probes, 1 kHz tone CS	$F_{4,907} = 1.33$ $p = 0.26$	$F_{1,907} = 82.5$ $p = 7 \cdot 10^{-16}$	$F_{4,907} = 1.53$ $p = 0.19$

Table 3.2 Results of two-way ANOVA test on eyelid PCs spike counts to different probe inputs on CR and nonCR trials.
Relevant to Fig. 3.5 C. Results are shown separately for three probe protocols used during recordings.

3.4.4 ROC analysis of Purkinje cells activity

As a further test that the cerebellar cortex computes binary choice we tested whether CR versus nonCR trials can be predicted from PC activity on a trial by trial basis. If binary choice is computed by the cerebellar cortex, these predictions should be equally strong across different probes that give rise to different CR probabilities. To test this hypothesis I employed ROC analysis (Britten et al., n.d.; Liu et al., 2012), a tool commonly used in decision-making studies. Spike count distributions from eyelid PCs on CR (cyan and violet) and nonCR (black) trials are shown in Fig 3.6 A for two probe types. I calculated spike counts over 700 ms from CS onset. Based on these

distributions, I constructed ROC curves and calculated choice probabilities as the area under the ROC curve (AUC). Fig 3.6 B shows a summary for every probe type that produced sufficient number of CRs during recording sessions. Choice probabilities for every probe type were significantly above chance (permutation test, 5000 samples), with $p<0.001$ for 3 "middle" probes with CR probability near 50% for all three protocols. The ability to predict a behavioral response was not different for all probes within each protocol ($p>0.4$ for all pairwise comparisons in each protocol, permutation test, 5000 samples). This suggests that the relationship between behavioral decision and eyelid PCs activity stays the same and the cerebellum employs the same mechanism independent of CR probability observed for a given probe.

It is possible that the binary choice is evident in eyelid PC activity only because of a potential feedback to the cerebellum about the behavioral response. I verified that changes in eyelid PC activity precede onset of CRs by examining the choice probability value as a function of time. For that, PCs spike trains were divided into a series of 100 ms non-overlapping time-windows and choice probabilities were calculated over each time-window. Here, for each protocol, I combined trials from three "middle" probes, but identical results were obtained across the probe analysis (Fig. 3.7). With PC activity aligned to CS onset in Fig. 3.6 C, the probability to predict CR correctly rose above chance after 200 ms from CS onset and peaked near CS offset, paralleling proper timing of the PCs firing rate decrease on CR trials. When aligned to CR onset (Fig. 3.6 D), for all three protocols, choice probabilities showed an above chance value 100 ms prior to

CR onset ($p<0.001$, permutation test, 5000 samples). The prediction accuracy achieved the peak value (AUC = 0.87±0.04) at 200 ms after CR onset, corresponding to the time of peak eyelid velocity.

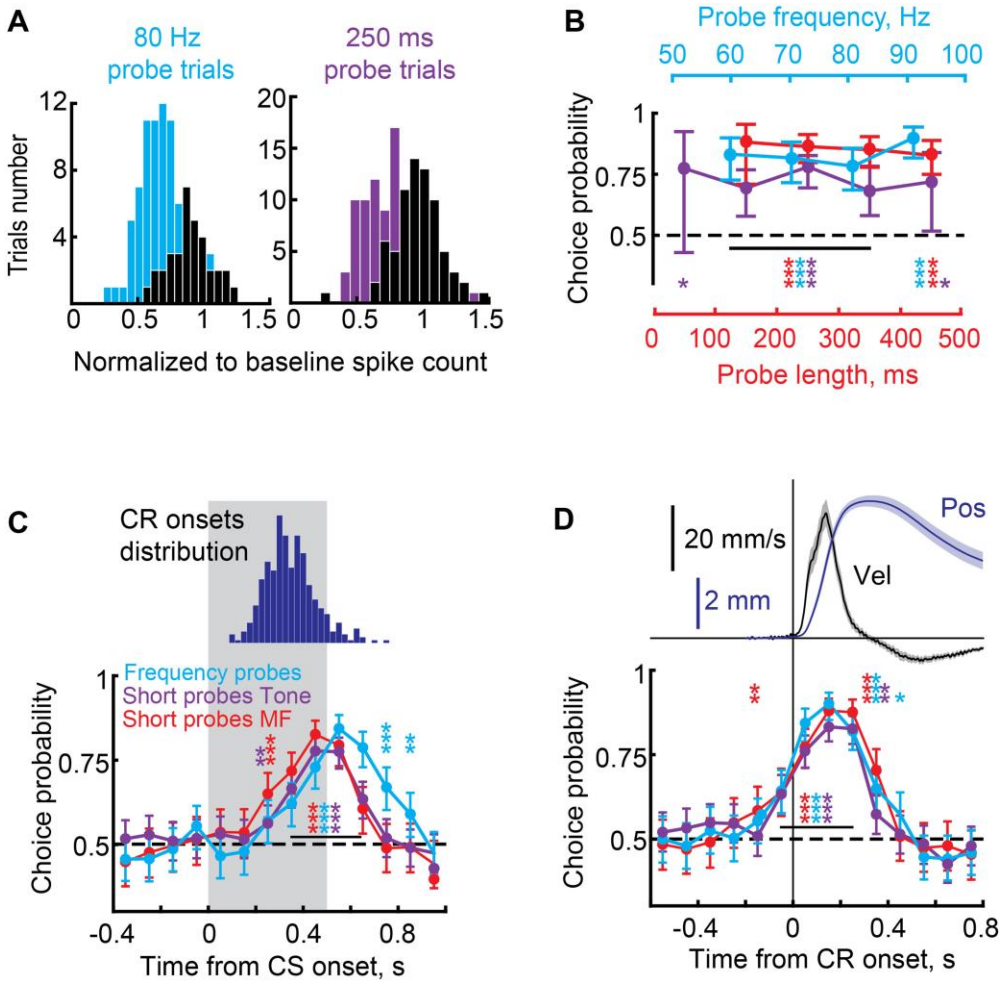

Fig. 3.6. ROC analysis of Purkinje cells activity.
A) Frequency distributions of PC spike counts on CR (cyan and violet) and non-CR trials (black). B) Choice probability calculated using ROC analysis for each probe type within three probe protocols. C) Choice probability as a function of time, using 100 ms time bins. Here, for each protocol we combined trials from three "middle" probes. Spike trains are aligned to CS onset. Points are plotted in the center of their corresponding bin. Distribution of CR onset times is shown on top. D) ROC analysis of the same data, but with trials aligned by behavioral CR onset (black vertical line).

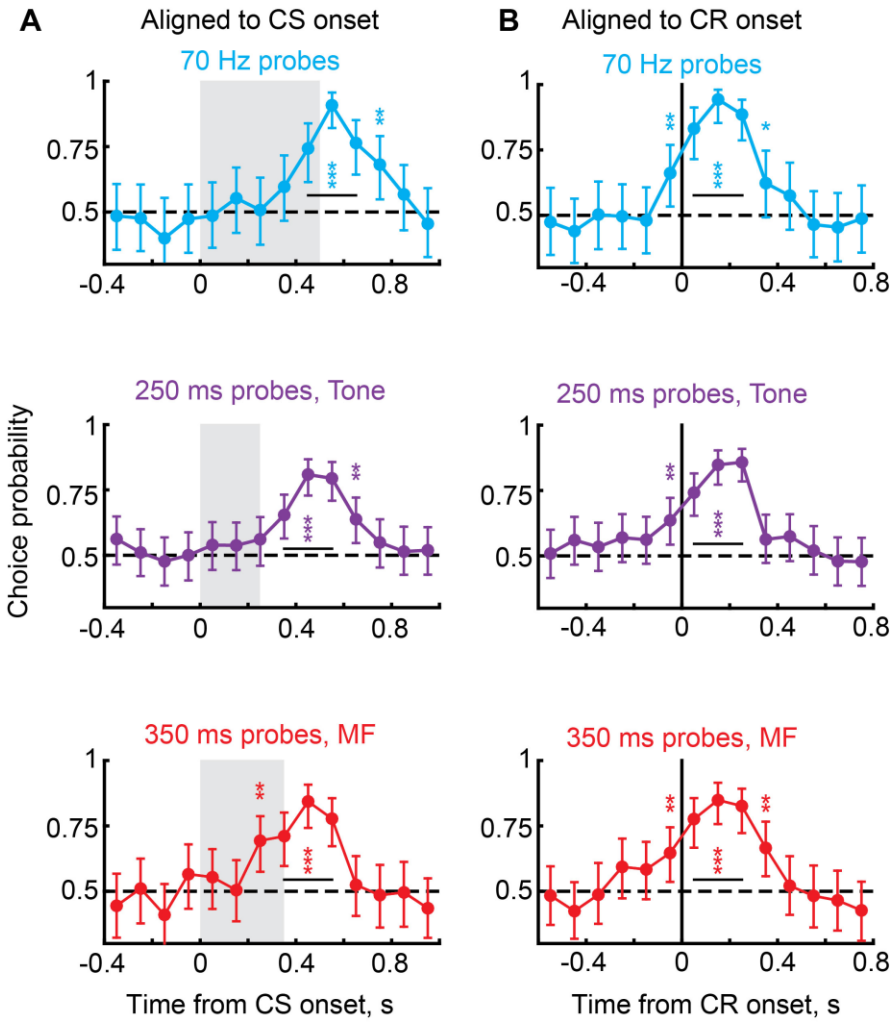

Fig. 3.7. ROC analysis of Purkinje cells activity on individual probe types.
A-B) Same as panels C-D of Fig. 3.6, but for individual probe types. A single example is shown per probe protocol: 70 Hz frequency probes (cyan), 250 ms short probes with tone as a CS (violet) and 350 ms short probes with mossy fiber stimulation as a CS (red). Region in grey indicates probe duration. In sum, results of ROC analysis for a single probe type parallel results obtained by combining trials from several probe types (panels C and D). For all panels error bars indicate 95% confidence intervals. Significance above chance at .05, .01 and .001 is depicted by 1, 2 or 3 stars respectively

3.4.5 Results of large scale cerebellar simulation suggest that DCN axon collaterals are necessary for binary choice

Next, I investigated possible network mechanisms underlying the phenomenon. A feedback mechanism can implement all-or-none decision on a circuit level (Wong and Wang, 2006). Binary choice, however, involves target-or-none responding. Thus, if feedback is a part of the phenomenon – it must participate in learning the target size of the response. Newly characterized axon collaterals from deep cerebellar nucleus neurons (DCNcol, Fig. 3.7 A) (Ankri et al., 2015; Houck and Person, 2015; Gao et al., 2016) synapse on granule and Golgi cells in cerebellar cortex as normal mossy fibers. By activating a subset of granule cells, which are a part of the well-established plasticity site in the cerebellum, DCNcol feedback can be used by the cerebellar cortex as an additional input to learn from. I tested the necessity of DCNcol for binary choice by using a large scale computer simulation of the cerebellum. The simulation is comprised of over 10^6 (2^{20}) spiking, conductance-based neurons and follows known connectivity rules of the cerebellum within the limits of the simulation size (see Methods). From the simulation we can monitor spiking activity of any (or all) neuron(s) along with virtual eyelid position calculated from simulation DCN output.

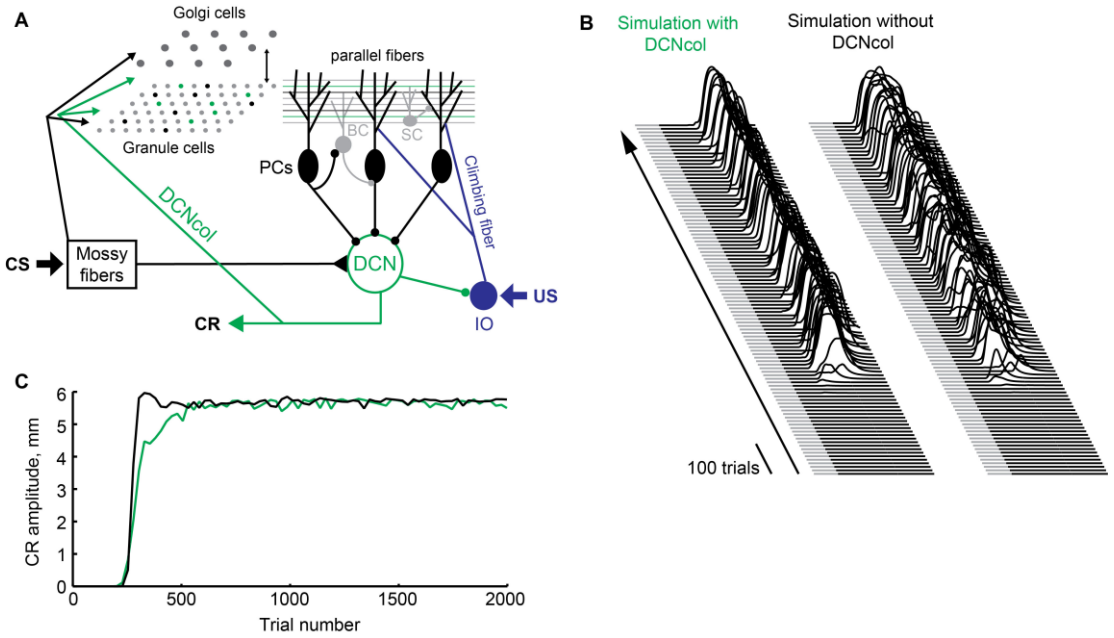

Fig. 3.8. Acquisition of CRs in large-scale cerebellar simulations.
A) Schematics of cell-types in the simulation and their connectivity. B) Waterfall plots of virtual eyelid position during acquisition of CRs from naïve state. Two simulations were identical except presence or absence of deep cerebellar nucleus axon collaterals (DCNcol) projecting back to the cerebellar cortex. Every 20th trial is plotted, training with Delay 500 ms protocol continued for 2000 trials. C) CR amplitude as a function of trial number, averaged across blocks of 10 trials. Green line corresponds to simulation with DCNcol present, black line – without.

We constructed two types of simulations differing only by the presence or absence of DCNcol (Fig. 3.8 A). Both simulations were trained using an ISI 500 ms protocol to produce full-sized CRs, each simulation acquired CRs at a similar rate (Fig. 3.8 B, C). After acquisition, we froze synaptic weights to prevent extinction and presented probe stimuli using the same three protocols used in the experiments. The summary of the simulation's behavioral output is shown in Fig. 3.9 A-C. For all

protocols, the simulation with DCNcol (Fig. 3.9 A) displayed binary choice: CR probability changed with the probes, but CR amplitude remained relatively constant, similar to the experimental data. Simulation without DCNcol, however, showed the breakdown of binary choice (Fig. 3.9 B) – CR amplitudes decreased as the probes were made increasingly different from the trained CS. To quantify the amount of independence of CR amplitude from CR probability I calculated a binarity index (BI, see Methods). By design, BI=1 corresponds to full independence, BI=0 corresponds to CR amplitude decreasing at the same rate as CR probability and BI<0 corresponds to a more abrupt decrease in CR amplitude than probability (Fig. 3.10). Both experimental data and data from simulation with DCNcol showed BI close to 1, while simulation lacking DCNcol had near zero or negative BI. Using the same method, I investigated the binarity of eyelid PCs responses. Similarly to the behavioral data, the responses of real PCs and virtual PCs from simulations with DCNcol were largely the same on CR trials, independent of probe and CR probability (Fig. 3.9 D, F). In contrast, in simulations without DCNcol, PC responses gradually changed with CR probability (Fig. 3.9 E, F).

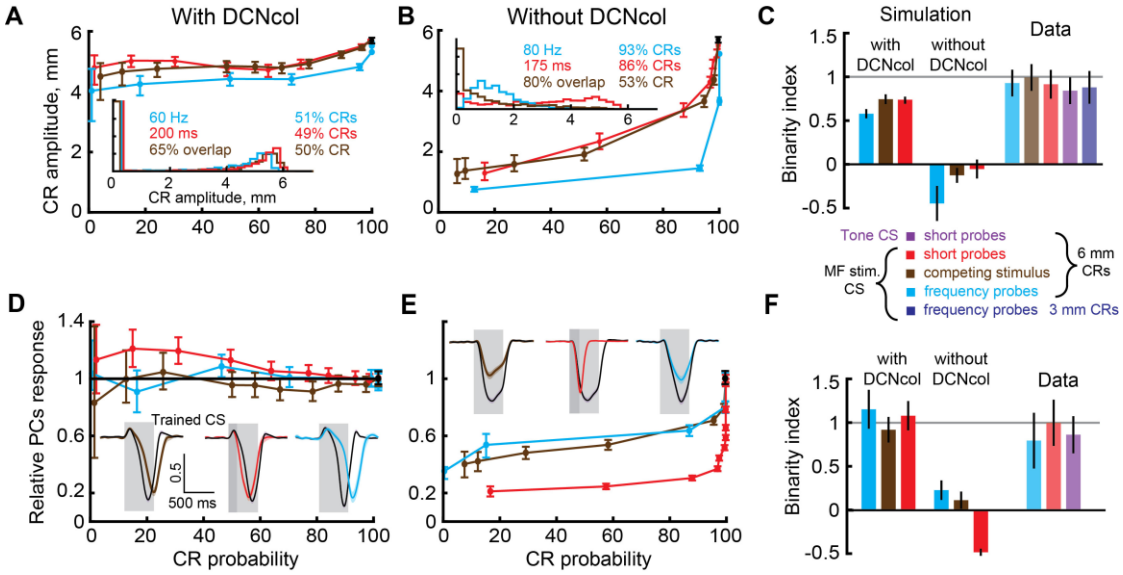

Fig. 3.9. Simulation with DCNcol replicates binary choice phenomenon.
A) Dependence of CR amplitude on CR probability in simulation with DCNcol. Results are shown for the same three probe protocols used in experiments. Color-coding is preserved as in Fig. 1 and further explained in legend on the right. Inset: distribution of CR amplitudes in response to one probe type per protocol, indicated above distribution. Error bars indicate 95% confidence intervals. B) Results of the same experiments performed in simulation without DCNcol. C) Comparison of behavioral results using binarity index (BI). Results are shown for two types of simulations and experimental data (lighter bars). Error bars show one standard deviation. D) Decrease in PCs activity on probe trials with CRs as a function CR probability. Insets: average normalized firing rate of PCs on CR trials, in response to trained CS (black) and probes (colored). E) Same as D, but for simulation without DCNcol. F) Same metric and layout as in C, but here eyelid PCs activity was used to determine the degree of binarity of PCs responses.

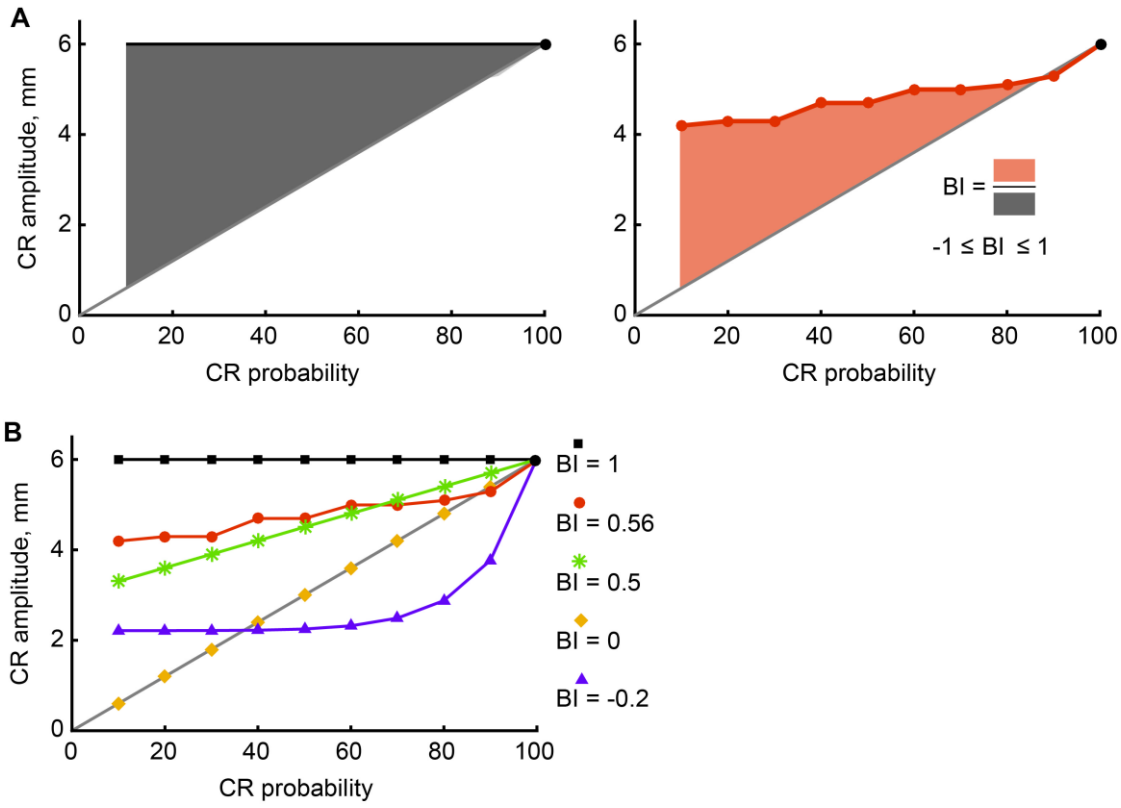

Fig. 3.10. Definition of binarity index.
A) The binarity index (BI) is designed to quantify the amount of independence between CR probability and response measure (CR amplitude or PC response). Black dot represents CR amplitude and probability in response to trained CS. The grey diagonal line is drawn such that CR amplitude decreases at the same rate as CR probability. Black line on the left panel represents an ideal case of CR amplitude being fully independent from CR probability. Red line on the right panel illustrates a possible shape of the experimental curve. BI is calculated as ratio between red and black shaded areas. The value of BI is defined between $-1 \leq BI \leq 1$. B) Five examples of possible dependence of CR amplitude on CR probability. BI = 1 corresponds to a full independence of CR amplitude from CR probability (black squares). A positive BI, but lower than one corresponds to a lower rate of decrease in CR amplitude than in CR probability (red circles and green asterisks). BI = 0 implies a proportional scaling of CR amplitude with probability (yellow diamonds). Finally, a negative BI indicates a faster decrease in CR amplitude than in CR probability (violet triangles).

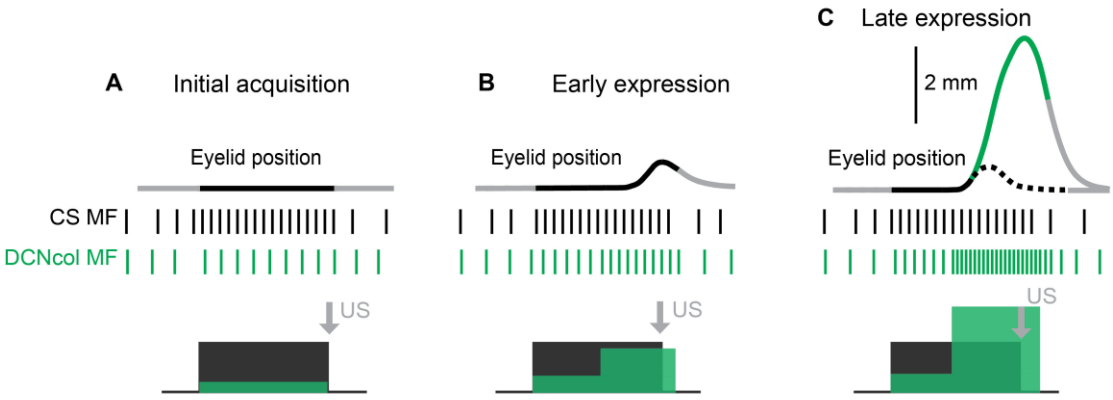

Fig. 3.11 Mechanism of possible DCN axon collaterals contribution to the binary choice phenomenon.
The layout of all panels is the following: eyelid position as a function of time is shown on top, with a region colored in black (or green) indicating the duration of CS. Vertical lines in the middle represent spike-train inputs from CS activated mossy fibers (black) and DCNcol mossy fibers (green) to the cerebellar cortex. Correspondingly colored rectangles below show schematized PSTHs. The time of US delivery is shown by a black arrow. A) Initial acquisition. Since CRs have not developed yet, CS activated mossy fibers are the main input that the cerebellum is learning to. B) Early expression. As small amplitude CRs start to happen, the cerebellum starts to get two kinds of inputs: 1) mossy fiber activated by CS (black) and 2) DCNcol mossy fibers (green). Initially, small CRs are associated only with CS, since that was the only input present during initial acquisition. C) Late expression. As CR amplitude increases, so does DCN activity that drives CR and therefore also DCNcol input (green) to the cerebellar cortex. After the acquisition is complete, only the early porting of CR is associated with CS (black portion of eyelid position profile), while the later portion of CR, determining its amplitude (green portion), is associated with DCNcol feedback input. Dotted black line represents a schematized CR profile that would be present if DCNcol feedback was disabled.

3.5. Discussion

In this project I utilized advantages of the eyelid conditioning paradigm to investigate possible cerebellar adaptations to noise and inputs ambiguity caused by it. Several studies have addressed this question in some way using different cerebellar-dependent paradigms: vestibular-ocular-reflex learning (VOR) (Guo et al., 2014), smooth pursuit (Medina and Lisberger, 2007) or eyelid conditioning (Svensson et al., 2010; Jirenhed and Hesslow, 2011). In (Guo et al., 2014) and (Medina and Lisberger, 2007) a visual stimulus was used in behavioral paradigms, which makes it impossible to explicitly disassociate the cerebellar contribution from transformation done by the visual system upstream from the cerebellum. To overcome this limitation I used direct electrical stimulation of mossy fibers as the trained CS and probe inputs to the cerebellum. While this method of delivering input to the cerebellum has been often utilized in eyelid conditioning, the majority of studies used stimulation with a fixed frequency for both acquisition and testing the expression of CRs (Steinmetz et al., 1985, 1986; Kreider and Mauk, 2010; Kalmbach et al., 2011). A series of papers from the Hesslow group have explored how the activity of PCs change when the mossy fiber stimulation CS is only partially presented (Jirenhed and Hesslow, 2011) or the frequency of stimulation is increased (Svensson et al., 2010) compared to the trained CS during acquisition. In these studies, however, there was no behavioral expression of the learned response due to the experimental design and the single unit data along with analysis were

limited. Interestingly though, one can spot 'binary choice' in raster plots of PCs responses to short probes shown in (Jirenhed and Hesslow, 2011). The lack of behavioral responses in their design unfortunately made it impossible to infer whether PCs responses became more variable or whether cerebellar output and the corresponding behavioral responses showed a probabilistic binary choice.

Cerebellar simulations revealed that DCN axon collaterals do not act merely as a positive feedback loop as suggested previously (Gao et al., 2016). Rather, the feedback about a response serves as an additional signal that contributes to learning the later portion of a response (Fig. 3.11). In this way, CR onset is associated with the sensory CS, while the later portion of the CR where the peak amplitude is achieved is associated with DCNcol feedback (Fig. 3.11). Because of this, noisy inputs that would have otherwise produced a range of CR amplitudes need to only initiate a response. Once a CR begins, activation of the DCNcol feedback provides the same signal back to the cerebellar cortex that had already been learned and CR amplitude is completed to the target amplitude.

The computational principle behind the binary choice can be used by several brain regions, such as hippocampus. In the cerebellum, the granule cell layer has long been implicated in pattern separation, a process that recodes similar mossy fiber inputs into less similar patterns of granule cell activity, presumably in the service of better discrimination (Marr, 1969; Billings et al., 2014). Our results reveal a form of pattern –

or in this case – amplitude completion, where DCNcol feedback ensures that response amplitudes maintain adaptive value. A parallel sequence of processing is seen in hippocampus where the granule cells in the dentate gyrus are thought to implement pattern separation while the next stage of processing in the CA3 region is thought to implement pattern completion (O'Reilly and McClelland, 1994). The sequence of pattern separation followed by pattern completion can help to discriminate between similar inputs, while ensuring that a proper "completed" response is produced if the similarity is sufficient. These apparent parallels may reflect somewhat general mechanisms for minimizing the detrimental influence of noise on performance of brain systems.

CHAPTER 4

Cerebellar implementation of movement sequences

4.1 Abstract

The cerebellum has been long implicated in learning and execution of motor skills (Krupa et al., 1993; Lisberger, 1994; Grafton et al., 2002; Shin and Ivry, 2003; Yttri and Dudman, 2016). The majority of movements we perform are not singular but are composed of a sequence of movements. Patients with cerebellar lesions are known to display either severe deficits (Doyon et al., 1997; Shimansky et al., 1997) in sequence learning, or are not able to learn the sequence at all(Shin and Ivry, 2003). While these studies point to the involvement of the cerebellum in learning a movement sequence, how this involvement is implemented is not known. Eyelid conditioning provides a cerebellar-dependent behavior (Krupa et al., 1993) and the ability to restrict and control inputs to the cerebellum though electrical stimulation of mossy fibers (Steinmetz et al., 1985). Using these advantages, I designed a training protocol to explicitly test the ability of the cerebellum to chain together a series of movements through associative learning processes. Our results demonstrate a simple yet general framework for how the cerebellum can use learning to chain together appropriately timed responses to produce a movement sequence.

4.2 introduction

Sequences of movements are a ubiquitous part of our lives. From a complex routine that a gymnast performs at the parallel bars to dancing or speech production - we learn skilled sequences through extensive practice and exercise. Both the cerebellum and basal ganglia have been long implicated in learning and execution of motor skills (Krupa et al., 1993; Lisberger, 1994; Doyon et al., 1997, 2002; Hazeltine and Ivry, 2002; Shin and Ivry, 2003; Lehéricy et al., 2005; Seidler et al., 2005; Yttri and Dudman, 2016). However, their relative involvement in learning and producing a motor sequence might differ.

While several imaging (Doyon et al., 2002; Lehéricy et al., 2005; Seidler et al., 2005) and recording (Jog et al., 1999) studies have reported changes in basal ganglia activity during sequence learning, studies involving lesions or neurodegenerative diseases were generally less conclusive. Studies in Parkinson's patients showed only a partial (Jackson et al., 1995; Doyon et al., 1997; Shin and Ivry, 2003) impairment of sequence learning compared to control groups and patients with focal basal ganglia lesions did not display deficits (Shin et al., 2005). Moreover, pharmacological lesions of globus pallidus internus, the primary output region of basal ganglia, resulted in a decrease of movement velocity and acceleration, but did not impair the production of learned sequences of movements (Desmurget and Turner, 2010). On the other hand, movement sequences and multi-joint movements are particularly sensitive to cerebellar dysfunction. For example,

one of the hallmark deficits of cerebellar pathology is dysdiadochokinesia (Diener and Dichgans, 1992) – an inability to perform a rapid alternating sequence of movements. Patients with cerebellar lesions display either severe deficits (Doyon et al., 1997; Shimansky et al., 1997) in sequence learning or are not able to learn the sequence at all (Shin and Ivry, 2003), although impaired learning was also present for directly cued movements (Spencer and Ivry, 2009). Overall, these studies point to the involvement of the cerebellum in learning a movement sequence.

The classical property of cerebellar learning is the ability to learn a predictive response (Marr, 1969; Bastian, 2006; Shadmehr et al., 2010; Therrien and Bastian, 2015). Several studies (Desmurget and Turner, 2010; Rünger et al., 2013) have shown that with repeated training, movement components in the sequence start to be initiated predictively, indirectly implying a strong cerebellar contribution. As an experimental paradigm, eyelid conditioning provides a cerebellar-dependent behavior (Krupa et al., 1993) where responses are initiated predictively, and provides the ability to restrict and control inputs to the cerebellum though electrical stimulation of mossy fibers (Steinmetz et al., 1985). Using these advantages, we designed a training protocol to explicitly test the ability of the cerebellum to chain together a series of movements through associative learning processes. Our results demonstrate a simple yet general framework for how the cerebellum can use learning to chain together appropriately timed responses to produce a movement sequence.

4.3 Methods

4.3.1 Initial training.

For initial training, subjects were given daily eyelid conditioning sessions comprised of 12 blocks of 9 trials each. All subjects were initially trained at an inter-stimulus interval (ISI) of 500 ms to produce left eye CRs. Each block consisted of 1 CS-alone trial and 8 paired trials. After reaching a robust CR performance, subjects were switched to the sequence training protocols.

4.3.2 Sequence training protocols.

4.3.2.1 Ipsilateral sequence of CRs.

Training sessions were composed of 12 (or 8 in rare cases) blocks of 9 trials each. Each block consisted of 1 CS-alone trial and 8 paired trials. On paired training trials subjects were presented with the same length of mossy fiber stimulation as the CS (500 ms) used during initial training, but with the US presented at one of two different times: the first time, designated as US_1, was at CS offset as with normal training. The second time, designated as US_2, occurred at 600 ms after CS offset. The factor that determined whether the US was presented at US_1 versus US_2 was the amplitude of the CR elicited by the mossy fiber stimulation CS. On trials when CR amplitude was lower than the target (3mm, corresponding to half-sized CR), US_1 was presented. The purpose of US_1 was to maintain robust responding of the first CR. If amplitude of the first CR was higher than

the target, US_1 was omitted and US_2 was presented. Since US_1 and US_2 were both delivered to the left eye, their intensity was the same. Some parameters of the US (pulse width and intensity) were adjusted to decrease the amount of eyelid squinting while still being sufficiently strong to support learning and expressions of CRs.

4.3.2.2 Contralateral sequence of CRs.

Training sessions were composed of 12 (or 8 in rare cases) blocks of 9 trials each. Each block consisted of 1 CS-alone trial and 8 paired trials. On paired training trials subjects were presented with the same length of mossy fiber stimulation as the CS (500 ms) used during initial training, US delivery was automatically determined by the following rule: if left eye CR amplitude was lower than the target (3mm, half-sized CR), US_L was presented to the left eye to maintain robust responding of the left eye CR. If amplitude of the left CR was higher than the target, US_L was omitted and US_R was presented to the right eye. During initial acquisition, the interval between CS offset and US_R was typically 400 ms (N=4), but for some subjects was 300 ms (N=1) or 500 ms (N=1). I chose to use a shorter duration of the gap interval relative to the ipsilateral sequence training, as our pilot data showed that most subjects were unable to learn a contralateral sequence with 600 ms gap interval from the naïve right eye state. After acquisition of the contralateral sequence of CRs subjects were switched to a 500ms gap interval training if shorter gap duration was used initially.

4.3.3 Extinction of following responses in the sequence.

The modification to the session design from initial training was straightforward. First, the left eye CRs elicited by the mossy fiber stimulation CS were still reinforced with US_1 (US_L) if CR amplitude was smaller than 3mm. However on trials with the left eye CR larger than 3mm US_2 (US_R) was no longer delivered compared to the acquisition training sessions.

4.3.4 Extinction of the first CR in the sequence while reinforcing the following CR.

Results from sessions described here are shown in Fig. 5. On paired trials I stopped delivering US_1 (US_L) regardless of the first left eye CR amplitude, while always delivering US_2 (US_R). The purpose of US_1 (US_L) was to reinforce performance of the first CR, hence its absence lead to extinction of the first left eye CRs. If the following responses in the sequence were driven by the mossy fiber stimulation CS, they should remain due to the reinforcing US_2 (US_R) delivery. Extinction of first left eye CRs was faster in the contralateral than ipsilateral protocol, as a consequence, subjects in ipsilateral sequence protocol were given at least two consecutive extinction sessions, while for subjects trained in the contralateral protocol one session was sufficient in most cases.

4.3.5 CS2 test sessions.

Subjects were trained in the sequence of CRs protocols using a mossy fiber stimulation CS, referred to in this section as CS1. In addition, subjects were trained at ISI 500 to produce left eye CRs with CS2. CS2 was either also a 500 ms long electrical stimulation of mossy fibers delivered through a separate electrode (N = 3 and N = 2 for subjects trained in the ipsilateral or contralateral sequence of CRs) or a 500 ms 1 kHz tone (N = 2 and N = 3 respectively). After robust responding to CS2 was reached, subjects were switched to CS2 test sessions. CS2 was never used for sequence training, nor for prior training of the left eye CRs at ISIs other than 500 ms, nor was CS2 used for right eye CR training. Each session consisted of 108 trials. On 60% of the trials CS1 was delivered and sequence of CRs was reinforced with $US_1(US_L)$ or $US_2(US_R)$ as described in the training protocol. Another 25% of the trials were ISI 500 CS2 paired trials. The remaining 15% of trials were CS2-alone. I chose these ratios (and slightly adjusted them per subject if needed) to maintain a balance between: enough CS1 trials to keep performance of CRs in the sequence sufficient, enough paired CS2 trials to keep performance of CRs elicited by CS2 sufficient, and finally enough CS2-alone trials so that the number of CS2 test trials per session was sufficient.

4.3.6 In-vivo recordings and unit isolation.

Only well-isolated single units were used for analysis. I recorded 153 single units during the ipsilateral sequence of CRs training sessions. Out of those, 13 units were

classified as eyelid PCs and 28 units as non-eyelid PCs. All single unit analysis reported in this chapter was performed on simple spikes from eyelid PCs. Recordings where the unit was lost at any time during the session were not included in the analysis.

4.3.7 Eyelid data analysis.

All data analysis was performed using custom-written scripts in MATLAB. For each trial, 2,500 ms of eyelid position data (200 ms pre-CS, 2,300 ms post CS) were collected at 1kHz sampling rate and at 12 bit resolution. The data were stored to a computer disk for subsequent off-line analysis. Eyelid position data was passed through a low-pass Savitzky–Golay filter. Eyelid velocity was calculated as the derivative of eyelid position with a second-order accurate scheme, again passed through a low-pass Savitzky–Golay filter.

A small fraction of trials were discarded if: 1) upward eyelid movement exceeding the 0.3 mm CR criterion occurred during 200 ms before the CS onset or first 100 ms after CS onset, or 2) negative eyelid deviation below 0.5 mm at any time prior to US onset on paired trials and from the start of the trial to 500 ms after CS offset for CS-alone trials. As a result, 1.4% (242/17574) of the trials from sessions related to the ipsilateral sequence protocol and 5.3% (690/13109) of the trials from sessions related to the contralateral sequence protocol were discarded. A larger fraction of trials were discarded in the contralateral sequence of CRs sessions due to fluctuations of both left and right eyelid positions contributing to the exclusion criteria.

An eyelid response was counted as a conditioned response (CR) if CR amplitude reached the 0.3 mm criterion. The CR onset time was defined only for CR trials and was determined using a custom-written two-step algorithm. The first step was designed to detect the initial deflection of eyelid position away from the pre-CS baseline, while the second step used linear interpolation to determine the exact time of CR onset. CR latency to criterion was defined as the first time point when eyelid position deviated above CR criterion. Additional CR features, including time to peak of CR velocity, time to half of peak CR, time to 90% of CR peak and CR peak time were introduced and calculated to study the co-variation of timing measures of the CRs in the sequence.

Due to the complexity of possible eyelid response profiles compared to conventional eyelid conditioning sessions, CR amplitudes were calculated in the following way.

4.3.7.1 Ipsilateral sequence of CRs data.

First, trials were passed through the criterion design to exclude trials where subjects squinted after the first CR making it impossible to determine the presence or absence of the following CR. I calculated the difference between 1) the maximum eyelid position value between 400 ms and 1100 ms from CS onset and 2) the minimum eyelid position value between 800 ms and 1100 ms from CS onset (US_2 was delivered at 1100 ms from CS onset). The difference needed to exceed at least 1 mm for the trial to be included for further analysis. On CS-alone trials, the MATLAB function *findpeaks* with

MinPeakProminence = 0.5 was initially used to estimate the number of CRs on a given trial. For trials with the number of CRs equal to or larger than 2 and the first CR onset time is smaller than 500 ms (length of CS), the *findpeaks* function output was used to determine the number, peak times and corresponding amplitudes of the CRs. If the onset time of the first CR was larger than 500 ms, I made an estimate of whether the first CR was delayed or absent. Here the first CR amplitude was defined as the maximum eyelid position value between CS onset and 200 ms after CS offset; the second CR amplitude was defined as the maximum eyelid position value between 200 ms to 800 ms from CS offset. Similar analysis was done on paired trails. The peak time of the second CR was assumed to be within 100 ms of US_2. The onset time of the second (and later) CR was defined as the time of the minimum of eyelid position value between two consecutive CRs.

4.3.7.2 Contralateral sequence of CRs data.

For the contralateral sequence protocol, a correction was made to account for possible simultaneous movement of the right eyelid during a left eye CR. Though these instances were rare, occasionally a subject's right eyelid position would deviate from baseline during a left eye CR. If this type of deviation reached CR criterion, this movement could potentially be falsely identified as a right eye CR. In order to reduce the number of false positives, right eye CR amplitude was calculated by subtracting the

maximum of the right eyelid position during the CS (when the left eye CR would occur) from the overall maximum of the right eyelid position.

4.4 Results

4.4.1 The design of CRs sequence training protocols

In a traditional eyelid conditioning experiment activating mossy fiber inputs via presenting a sensory CS, or by direct electrical stimulation, is the cue the cerebellum uses to learn a predictive response. We tested the hypothesis that the cerebellum can learn to chain together a sequence of movements by using feedback signal(s) (FS) from one component of a movement to serve as the "CS" for the next component in the sequence. Thus, I designed experiments such that only the first component of movement can be learned thorough association with an external sensory stimulus. The delivery of an external stimulus was accomplished via direct electrical stimulation through electrodes implanted in the middle cerebellar peduncle, a structure that is solely comprised of mossy fibers projecting to the cerebellum. As such, all experiments employed direct electrical stimulation of mossy fibers as the conditioned stimulus (CS) to train what would become the first response of a sequence. These procedures ensured that the CS was restricted to the cerebellum and did not propagate to areas that could, in principle, provide a delayed

secondary input to the cerebellum (Halverson et al., 2010; Siegel and Mauk, 2013). Initially subjects were trained at ISI 500 ms.

4.4.2 Training ipsilateral sequence of CRs

We began by employing a training protocol designed to ask whether the expression of an eyelid CR elicited by mossy fiber stimulation as the CS can itself serve as the CS for a CR that occurs later. Once subjects reached robust responding to the mossy fiber CS, each rabbit was switched to the CR sequence training protocols. The design of the ipsilateral sequence training protocol is illustrated in Fig. 4.A. On paired training trials subjects were presented with the same mossy fiber stimulation CS (500 ms) used during initial training, but now with the US presented at one of two different times: the first time, designated as US_1, was at CS offset as with normal training. The second time, designated as US_2, occurred at 600 ms after CS offset. The factor that determined whether the US was presented at US_1 versus US_2 was the amplitude of the CR elicited by the mossy fiber stimulation CS. On trials when CR amplitude was lower than the target (3mm, half-sized CR), US_1 was presented. The purpose of US_1 was to maintain robust responding of the first CR. If amplitude of the first CR was higher than the target, US_1 was omitted and US_2 was presented. Previous studies (Kalmbach et al., 2009) established that the temporal gap larger than 400 ms between the offset of a mossy fiber stimulation CS and US onset does not support learning of eyelid CRs. Since in the

sequence training protocol the gap between CS offset and US_2 was larger than 400 ms, mossy fiber stimulation CS was not sufficient to support learning of the second CR timed to US_2. Therefore, if subjects were able to learn the second CR corresponding to the time

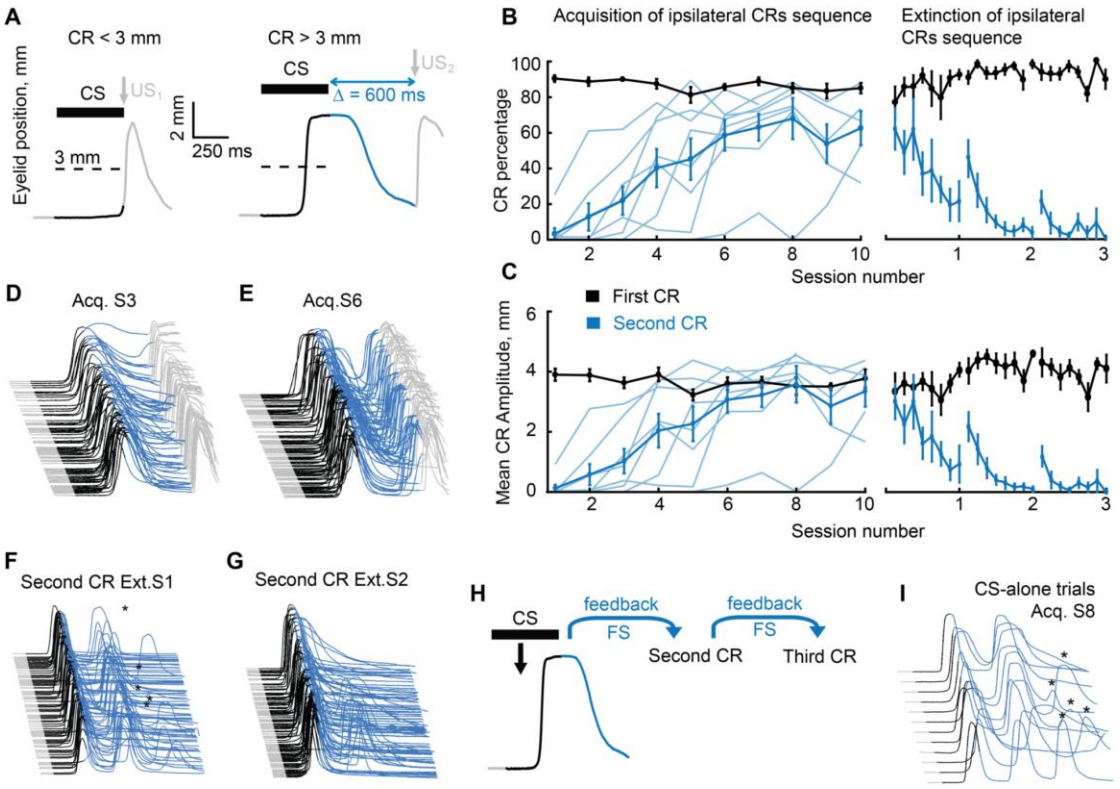

Figure 4.1. Acquisition and extinction of ipsilateral sequence of CRs.
A) Schematics of ipsilateral sequence training protocol. Black color indicates CS duration, blue – gap interval between CS offset and US_2 onset, grey – periods before CS onset and after US onset. Left panel shows example trial with first CR amplitude smaller than target 3 mm value (indicated by a dotted line). Thus, by design of training procedure US_1 is delivered. Right panel shows a trial with first CR amplitude larger than 3 mm. Thus, US_1 is omitted and US_2 is delivered. B) CR probability as a function of session number. Left panel shows acquisition curves of second CR in ipsilateral sequence. Probability of the first CR is shown in black, the second – in blue. Thin lines represent individual subjects, thick lines – across subjects' averages. Right panel shows data from three sessions of second CR extinction. Data in each session was broken down into eight equal portions to evaluate the time profile of extinction through the session. C) Same as B), but for average amplitude of CRs in ipsilateral sequence. (D-G) Example acquisition and extinction sessions of ipsilateral sequence of CRs. In all cases only trials with first CR amplitude larger than 3 mm are shown, arranged chronologically from bottom to top. Upward defilations indicate closure of the eyelid. H) Illustration showing why subjects trained at ipsilateral sequence of CRs should produce a third CR on CS-alone trials without it being explicitly reinforced during training. I) Example CS-alone trials from late acquisition session. The third CRs in a sequence are indicated by asterisks above.

of US_2, that alone would suggest that the cerebellum uses the feedback information about the first CR as a new "CS" to learn the subsequent CR.

All subjects successfully acquired a sequence of ipsilateral CRs (Fig. 4.1 B-E) with the timing of the second peak appropriate for the time at which US_2 was presented. The probability and amplitude of the second CR (Fig. 4.1 B and C respectively, blue lines) grew monotonically over several sessions of training and eventually reached asymptotic value, paralleling a typical acquisition curve in eyelid conditioning ($F_{(9,63)}$ = 6.44, $P = 2 \cdot 10^{-6}$; $F_{(9,63)}$ = 7.16, $P = 5 \cdot 10^{-7}$ for second CR probability and amplitude respectively). To test whether the second CRs were acquired through associative learning, subjects underwent three extinction sessions of the second CR (Fig. 4.1 B, C, F, G). As expected, the probability and amplitude of the second CRs (blue lines) monotonically decreased ($F_{(23,120)}$ = 4.96, $P = 3 \cdot 10^{-9}$; $F_{(23,120)}$ = 4.92, $P = 3 \cdot 10^{-9}$ for second CR probability and amplitude respectively) without any effect on the probability and amplitude of the first CRs (black lines) ($F_{(23,120)}$ = 1.45, P = 0.10, $F_{(23,120)}$ = 1.35, P = 0.15 for first CR probability and amplitude). Similar to extinction sessions in conventional eyelid conditioning protocols, I observed spontaneous recovery (Weidemann and Kehoe, 2004; Ohyama et al., 2010; Thanellou and Green, 2011) of the second CRs at the beginning of the second and third extinction sessions (two-tailed Student's t-test, $P < 0.03$ for both second CR probability and amplitude).

The primary hypothesis I sought to test here was that the cerebellum uses a FS of some sort about the first CR as a separate "CS" to learn the second CR (Fig. 4.1 H). Such FS should depend only on the expression of the first CR. Therefore, a prediction illustrated in Fig. 4.1 H can be made. When a well-trained subject is presented with a CS-alone trial, the mossy fiber stimulation CS will elicit the first CR, then FS from the first CR will elicit the second CR. If the second CR is large enough, the FS from it should be similar to the FS from the first CR. Thus, one would expect to see the third (and following) CRs on CS-alone trials, though subjects were never explicitly trained to produce them. This is indeed what I observed (Fig. 4.1 F, I). Third and following CRs appeared in late acquisition sessions with robust second CR performance and disappeared with the extinction of the second CRs (Fig. 4.1 F, G). Due to the technical limitations (see methods), it was possible to monitor up to four ipsilateral CRs.

4.4.3 Training contralateral sequence of CRs

Complex movements typically involve more than one muscle group and often also bilateral coordination (Kelso et al., 1979; Castiello et al., 1993). I therefore asked whether this training protocol could be implemented to produce a sequence of contralateral eyelid CRs (Fig. 4.2 A). As before, subjects were initially trained with a mossy fiber stimulation CS to produce left eye CRs at ISI 500 ms. After successful acquisition, subjects were switched to a contralateral CRs sequence protocol. Here, if the left eye CR amplitude was lower than the target (3mm, half-sized CR), US_L was

presented to the left eye to maintain robust responding of the left eye CR. If amplitude of the left CR was higher than the target, US_L was omitted and US_R was presented to the right eye. During initial acquisition, the interval between CS offset and US_R was typically 400 ms (N=4), but for some subjects was 300 ms (N=1) or 500 ms (N=1). I chose to use a shorter duration gap interval relative to the ipsilateral sequence training, due to the pilot data showing that most subjects were unable to learn a contralateral sequence with a 600 ms gap interval from the naïve right eye state. In this situation, however, there is less concern about the ability of the mossy fiber stimulation CS to drive the second CR, since the mossy fiber stimulation CS was delivered through electrodes implanted in the left middle cerebellar peduncle. The only way that CS could propagate to the right cerebellar hemisphere was by antidromic activation of neurons in the pontine nucleus that have bilateral axon projections. Since the number of such neurons is extremely low (Tan and Gerrits, 1992; Serapide et al., 2002; Kratochwil et al., 2017), it is unlikely that unilateral mossy fiber stimulation CS could support acquisition of right eye CRs. Several control experiments described later in the paper confirm this.

All subjects successfully acquired the contralateral sequence of CRs (Fig. 4.2 B-E, $F_{(9,46)} = 14.28$, $P = 1 \cdot 10^{-10}$; $F_{(9,46)} = 9.05$, $P = 1 \cdot 10^{-7}$ for right CR probability and amplitude respectively). From the design of the contralateral sequence protocol one would not expect, unlike the ipsilateral eye protocol, to observe a third response to be present. Indeed, these responses were never observed. Contralateral acquisition was followed with three extinction sessions of right eye CRs (red lines in right portion of Fig.

2 B, C) which verified their associative nature ($F_{(23,96)} = 4.04$, $P = 7 \cdot 10^{-7}$; $F_{(23,96)} = 6.89$, $P = 5 \cdot 10^{-12}$ for right CR probability and amplitude respectively). Extinction of right eye CRs did not influence the performance of the preceding left eye CRs ($F_{(23,96)} = 0.96$, $P = 0.52$; $F_{(23,96)} = 0.82$, $P = 0.70$ for left eye CR probability and amplitude respectively). Spontaneous recovery of right eye CRs at the beginning of the second and third extinction sessions was present here as well (two-tailed Student's t-test, $P < 0.04$ for both right CR probability and amplitude).

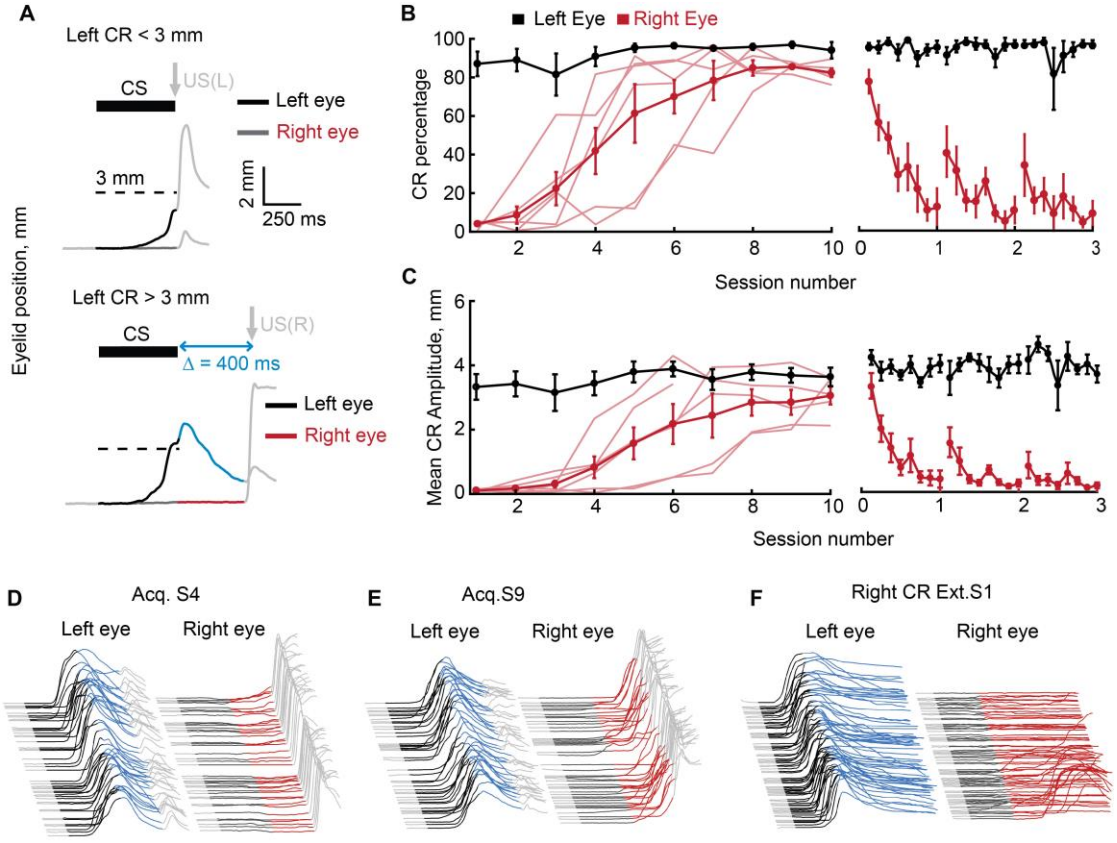

Figure 4.2. Acquisition and extinction of contralateral sequence of CRs.
A) Schematics of contralateral sequence training protocol. Left eyelid position is divided into a black portion, indicating CS duration, and a blue portion indicating the gap interval. Right eyelid position is shown in dark grey and red correspondingly. Light grey color indicates periods before CS onset and after US onset. Top panel shows example trial with left eye CR amplitude smaller than target 3 mm value (indicated by a dotted line). Thus, by design of training procedure USL is delivered to the left eye. Bottom panel shows a trial with left CR amplitude larger than 3 mm. Thus, US_L is omitted and US_R is delivered to the right eye. B) CR probability as a function of session number. Left panel shows acquisition curves of right eye CR in contralateral sequence. Probability of left eye CR is shown in black, right eye CR – in red. Thin lines represent individual subjects, thick lines – across subjects' averages. Right panel shows data from three sessions of right eye CR extinction. Data in each session was broken down into eight equal portions to evaluate the time profile of extinction through the session. C) Same as B), but for average amplitude of CRs in contralateral sequence. D-F) Examples of acquisition and extinction sessions of contralateral sequence of CRs. For each session left eye responses are shown on left, right eye responses – on right. In all cases only trials with left eye CR amplitude larger than 3 mm are shown.

4.4.4 Prediction 1: The first CR in the sequence is necessary for the following CR

If the cerebellum indeed generates the later CRs using a FS from the first CR, several testable predictions follow. The first prediction is straightforward: on trials without the first CR a subsequent second CR (ipsilateral protocol) or right eye CR (contralateral protocol) should not be present. To test this prediction I used eyelid responses on CS-alone trials from sessions with the second left eye CR probability > 40% (or right eye CR probability for contralateral sequence). Example trials from subjects trained at the ipsilateral or contralateral sequence of CRs are shown in Fig. 4.3. A, B. Data from all subjects on all CS-alone trials is shown in Fig. 4.3 C, D. For subjects trained in the ipsilateral sequence of CRs, the amplitude of the second CR plotted versus amplitude of the first CR is shown in Fig. 4.3. C. Each dot represents data from a single CS-alone trial. Similarly, Fig. 4.3 D shows right eye CR amplitudes versus left eye CR amplitudes for subjects trained to produce the contralateral sequence of CRs. Points corresponding to example trials are indicated in grey. Horizontal and vertical black lines represent nonCR cutoffs (CR amplitudes < 0.3 mm), diagonal is shown by a dashed black line. Fig. 4.3 E and F show the same data plotted as a probability of the second CR (or right eye CR) as a function of first CR amplitude.

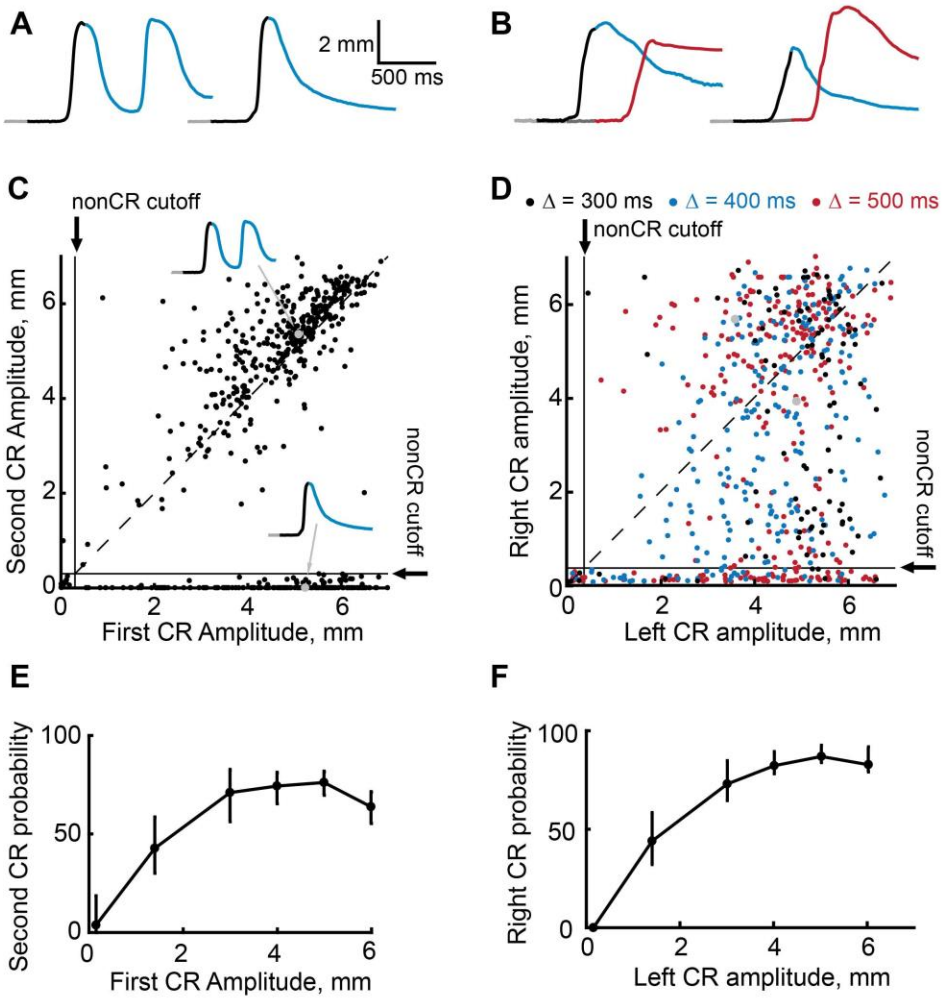

Figure 4.3. First CR is necessary for subsequent ones in the sequence.
A-B) Eyelid CRs on two example trials from ipsilateral and contralateral sequence training sessions respectively. Color-coding of time intervals is preserved from Figures 1 and 2. C) Amplitude of the second CR versus amplitude of the first CR in ipsilateral sequence. Each dot represents a single CS-alone trial. Vertical and horizontal black lines represent nonCRs cutoffs, dashed black line shows the diagonal. For panels C-D dots corresponding to example trials from panels A-B are shown in grey. D) Similar to C), but for subjects trained in contralateral sequence of CRs. Data obtained from sessions with different gap intervals is color-coded as indicated in legend. E) Average probability of the second CR as a function of first CR amplitude. Error-bars show 95% confidence intervals (obtained by bootstrapping with 0.3 mm nonCR threshold, 2000 samples). F) Same as E), but for subjects trained at contralateral sequence of CRs. Here we combined data from sessions with different gap intervals.

For both protocols, the probability of the second CR (or right eye CR) started to decrease on trials with the first CR amplitude smaller than 3 mm, which was used during sequence training (Chi-square analysis, $\chi^2(6, 571) = 70.13$, $P = 2 \cdot 10^{-13}$, $\chi^2(6, 578) = 92.4$, $P = 1 \cdot 10^{-21}$, for ipsilateral and contralateral sequence protocols respectively). Most importantly, on trials without the first CR there were no subsequent ones (Student's t-test by bootstrapping 10000 samples, P = 0.36 for ipsilateral sequence, P = 0.99 for contralateral sequence).

4.4.5 Prediction 2: Timing of CRs in sequence co-vary on trial-by-trial basis

A second prediction would be to observe trial-to-trial co-variability in timing of CRs in a sequence. Indeed, if FS from the first CR serves as a CS for the following CR, then on trials with early first CRs (relative to CS onset) should the following CRs also happen earlier than the average; on trials with late first CRs the following CRs should also happen later. Example trials with earlier (on top) and later (on bottom) first CR onset times are shown in Fig. 4.4 A and B from the ipsilateral and contralateral sequence respectively. Notice even from the example trials that the whole sequence of CRs is shifted with respect to the timing of the first CR. I investigated the degree of co-variation between the timing of CRs in sequence using a variety of CR timing measures, including CR onset time to CR peak time. These measures were defined for every CS-alone trial when the amplitude of both responses in a sequence was larger than 2 mm. For every

pairwise combination of CR timing measures we calculated a Pearson correlation coefficient. I found that the majority of combinations showed a significant trial-to-trial correlation for both ipsilateral and contralateral sequence of CRs ($P < 0.01$ for all shown pairs). Data from two such pairs is shown for each protocol in Fig. 4.4 C-F. For the ipsilateral sequence, while CR latencies to criterion showed significant trial-to-trial correlation (Fig 4.4 C), the largest correlation was observed between the peak time of the first CR and the early portion of the following CR (Fig. 4.4 E). For the contralateral sequence of CRs the timing of early to middle portions of CRs in sequence showed high correlations (Fig. 4 D, F). Overall, our results support the prediction of CR timing co-variability for both ipsilateral and contralateral protocols. Some differences between protocols in the most correlated portions of CRs could be attributed to several factors. First, I used longer gap intervals for the ipsi- versus contralateral sequence (600 ms versus 400-500 ms respectively). Second, for the ipsilateral sequence all responses were produced by the same eye and therefore the following response can start only after some minimal time required for the eye to open back from the previous CR. Together these results suggest that FS from the appropriate portions of the previous movement can be used by the cerebellum and depend on the time constraints imposed by the sequence training protocol.

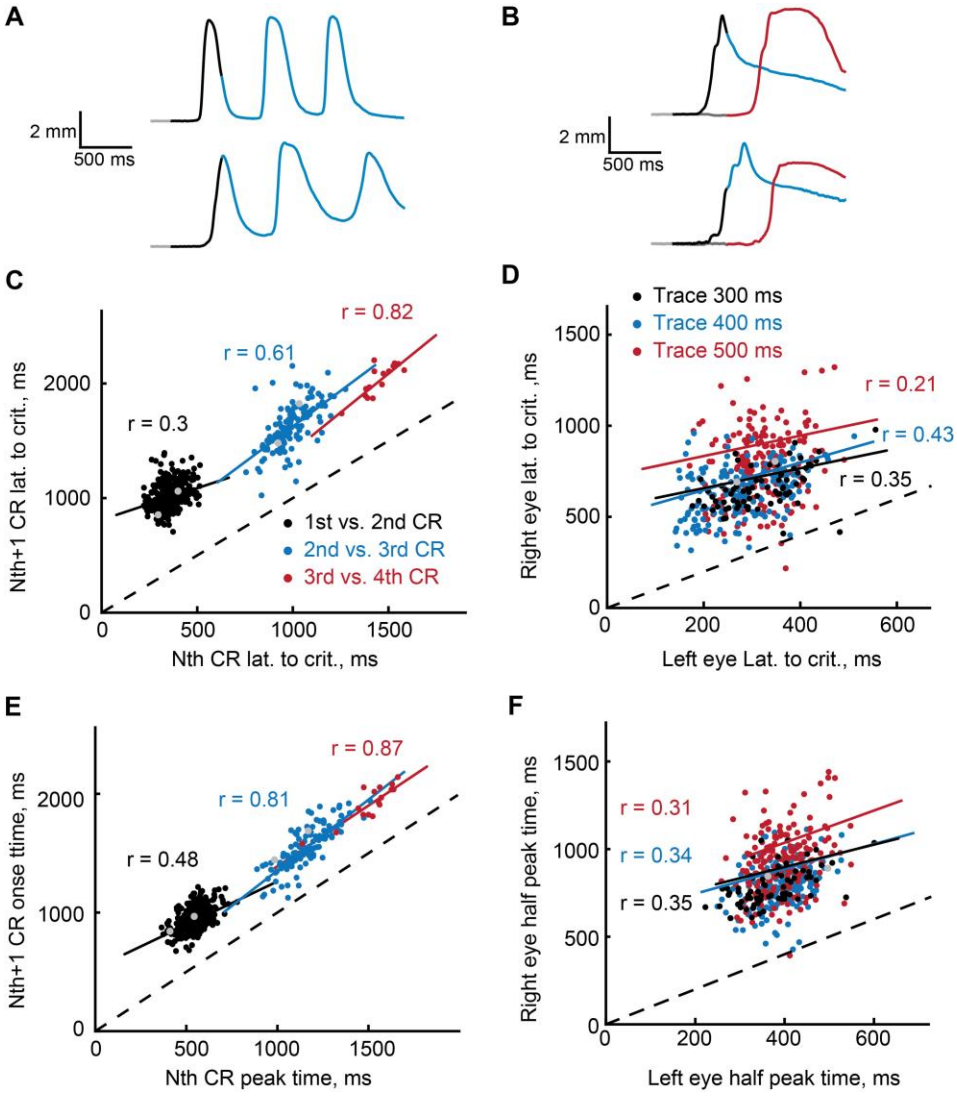

Figure 4.4. Timing of CRs in sequence co-varies from trial-to-trial.
A) Two example trials showing eyelid CRs from subjects trained in ipsilateral sequence of CRs. B) Similarly two example trials from subjects trained in contralateral sequence of CRs. C) Latency to criterion of following versus previous CRs in ipsilateral sequence. Each dot represents a single CS-alone trial, a dotted black line shows the diagonal. Colored lines show a linear regression fit to data from corresponding pair of CRs as indicated in the legend. For C-F dots corresponding to example trials from panels A and B are shown in grey. D) Similar plot for latency to criterion of CRs in contralateral sequence. Here colors indicate different gap intervals, as in the legend. E) Same for the peak versus onset time of CRs in ipsilateral sequence. F) Same for the time to half peak amplitude of CRs in contralateral sequence.

4.4.6 Prediction 3: following CR in a sequence is not elicited by mossy fiber stimulation CS

Finally, we designed two control experiments to test relatively directly whether the following CR in a sequence is associated with FS from the first CR and not with the mossy fiber stimulation CS itself. The setup of the first control experiment is shown in Fig. 4.5 A, D. Here I stopped reinforcing the left eye CRs with US_1 (US_L), while delivering US_2 (or US_R in contralateral sequence) at the same time as during the sequence training. Absence of US_1 should extinguish the first left eye CRs. Thus, if the following responses in a sequence will also extinguish despite the presence of reinforcing US_2 (or US_R), I would conclude that the FS from the first CR was used to learn the second CR. This is indeed what I observed. Fig. 4.5 B, E shows average CR amplitudes is a sequence as a function of block number. Since extinction typically took longer for subjects trained with the ipsilateral sequence, most subjects received two extinction sessions. As the amplitude of the left eye CR decreased, so did the amplitude of the second left eye CR in the ipsilateral sequence ($r = 0.86$, $P = 5 \cdot 10^{-13}$); or right eye CRs in the contralateral sequence ($r = 0.80$, $P = 1 \cdot 10^{-13}$).

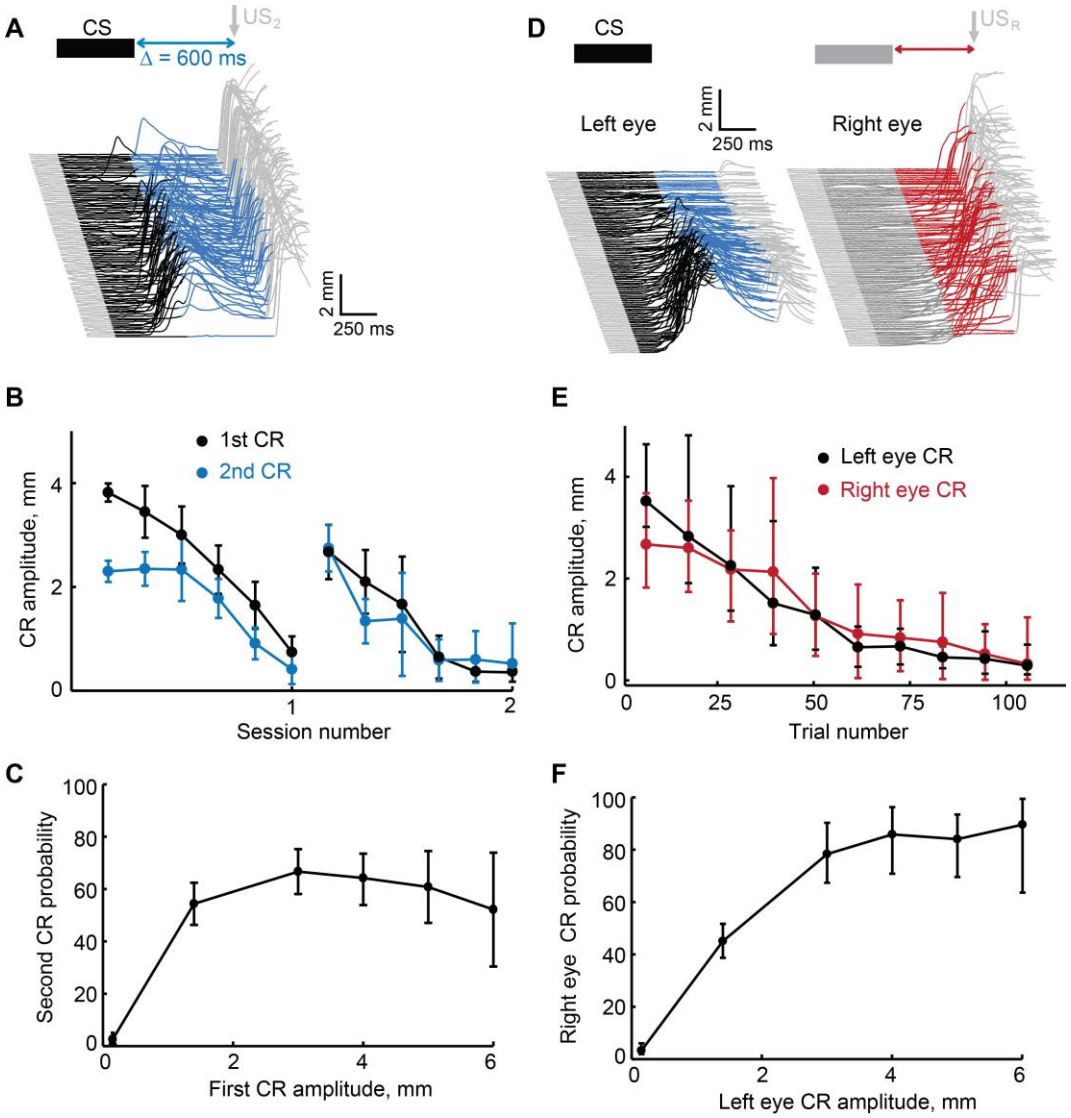

Figure 4.5. Extinction of the first CR in the sequence eliminates the following CRs.
A) Schematics of experiment setup and example session from subject trained in ipsilateral sequence of CRs. On all paired trials US_1 was never presented and US_1 was always presented, regardless of first CR amplitude. B) First and second CR amplitudes over two consecutive sessions of first CR extinction. Each point is an average across one sixth of trials in the session, error-bars represent standard error. C) Second CR probability as a function of first CR amplitude. Error-bars show 95% confidence intervals (obtained by bootstrapping with 0.3 mm nonCR threshold, 2000 samples). D-E) Similar analysis for subjects trained in contralateral sequence protocol.

In order to evaluate data at single trial resolution I performed analysis similar to the previously shown analysis in Fig. 4.3. Results are shown in Fig. 4.5 C, F. During normal expression of the CRs sequence, the probability of observing following responses decayed with first CR amplitude (Chi-square analysis, $\chi^2(6, 836) = 60.90$, $P = 2 \cdot 10^{-11}$, $\chi^2(6, 572) = 92.7$, $P = 4 \cdot 10^{-18}$, for ipsilateral and contralateral sequence protocols respectively). On trials without the first left eye CR, the probability of the following responses was negligible (second left eye CR probability = 2.6 ±0.9% for ipsilateral sequence protocol, right CR probability = 2.3 ±1.0% for contralateral sequence protocol).

If a FS is the cue that drives a second CR, then it should not be necessary for the first CR to be elicited by the original training CS. The schematic of a second control experiment that I used to test this prediction is shown in Fig. 6 A, E. I started by training subjects to produce left eye CRs (ISI 500 ms) with two different types of CS: CS1 and CS2. Electrical stimulation of mossy fibers was always used as CS1, CS2 was either a 500 ms mossy fibers stimulation delivered through a separate electrode (N = 3 and N = 2 for subjects trained in ipsilateral or contralateral sequence of CRs) or a 500 ms 1 kHz tone (N = 2 and N = 3 respectively). At the end of this pre-training each subject elicited robust CRs to the presentation of either CS1 or on separate trials to CS2. Then, during sequence training as described above, only CS1 was used, CS2 was never used for sequence training, nor for training left eye CRs at ISIs other than 500 ms, nor was it used for right eye CR training.

After successful acquisition of the ipsi- or contralateral sequence of CRs, subjects were then presented with occasional CS2-alone trials. This input should elicit a left eye CR, since this is how subjects were trained. However, if the following CRs in the sequence will also be present on CS2-alone trials, it would indicate that the cerebellum used FS from the first CR to learn the following CR in the sequence and not CS1. Again, this is indeed what I observed. Example eyelid responses on CS2-alone trials are shown in Fig. 4.6 B and Fig. 4.7 B for subjects trained with the ipsilateral or contralateral sequence respectively. On most CS2 alone trials when the first left eye CR was present I also observed the rest of the CRs in the sequence the subject was trained to with CS1. The summary across all sessions with CS2 test trials is shown in Fig. 4.6 C and Fig. 4.7 C. Here each dot shows the probability within a session of the second CR in a sequence, the color indicates a group based either on CS type (CS1 or CS2) or amplitude of the first left eye CR. On trials with the amplitude of first left eye CR larger than 3 mm, the probability of observing other responses in a sequence was the same on trials with either CS1 or CS2 (Tukey's post hoc test, $P = 0.73$, blue versus brown bars for ipsilateral sequence; $P = 0.51$; red versus brown bars for contralateral sequence). Importantly, on CS2 trials without the first response (violet bars), there were no other CRs in the sequence (ipsilateral sequence: two-sided Student t-test, $P = 0.35$, comparing brown versus violet bars – $F_{(3,61)} = 187$, $P = 1 \cdot 10^{-30}$; Tukey's post hoc test $P = 4 \cdot 10^{-9}$; contralateral sequence: two-sided Student t-test, $P = 0.34$, comparing brown versus violet bars – $F_{(3,62)} = 159$, $P = 5 \cdot 10^{-29}$; Tukey's post hoc test $P = 4 \cdot 10^{-9}$).

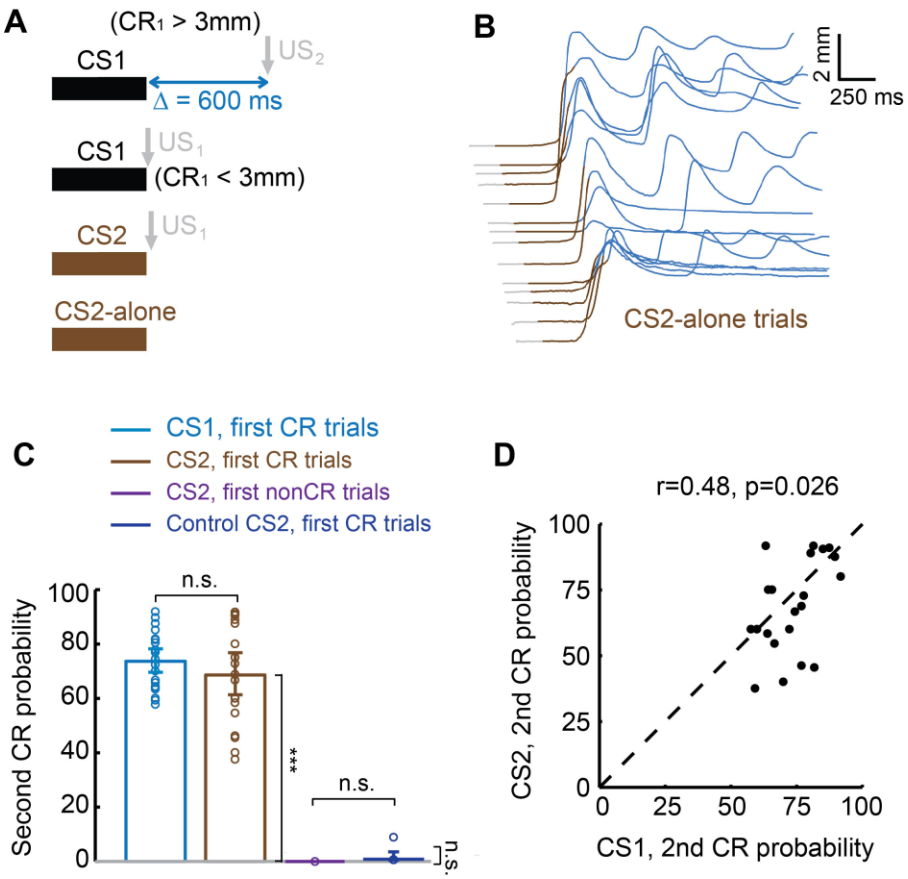

Figure 4.6. Sequence of CRs is present regardless of CS type that drives the first CR (ipsilateral sequence data).
A) Schematics of stimuli presentation during for types of trials in the session. CS1 was used for training ipsilateral sequence of CRs while CS2 was used only to train first CR. B) Example session showing eyelid responses on CS2-alone trials from subject trained in ipsilateral sequence with CS1. Brown color indicates CS2 duration. C) Second CR is present on either CS1 or CS2 trials only if first CR is present too. Each dot represents average second CR probability over corresponding trial type across one session, bars show a global average across all CS2 test session. Error bars show 95% confidence intervals (obtained by bootstrapping with 0.3 mm nonCR threshold, 2000 samples). D) Second CR probability on CS2 versus CS1 trials co-varies between sessions. Each dot represents an average across corresponding trial type during one session, dashed line shows the diagonal.

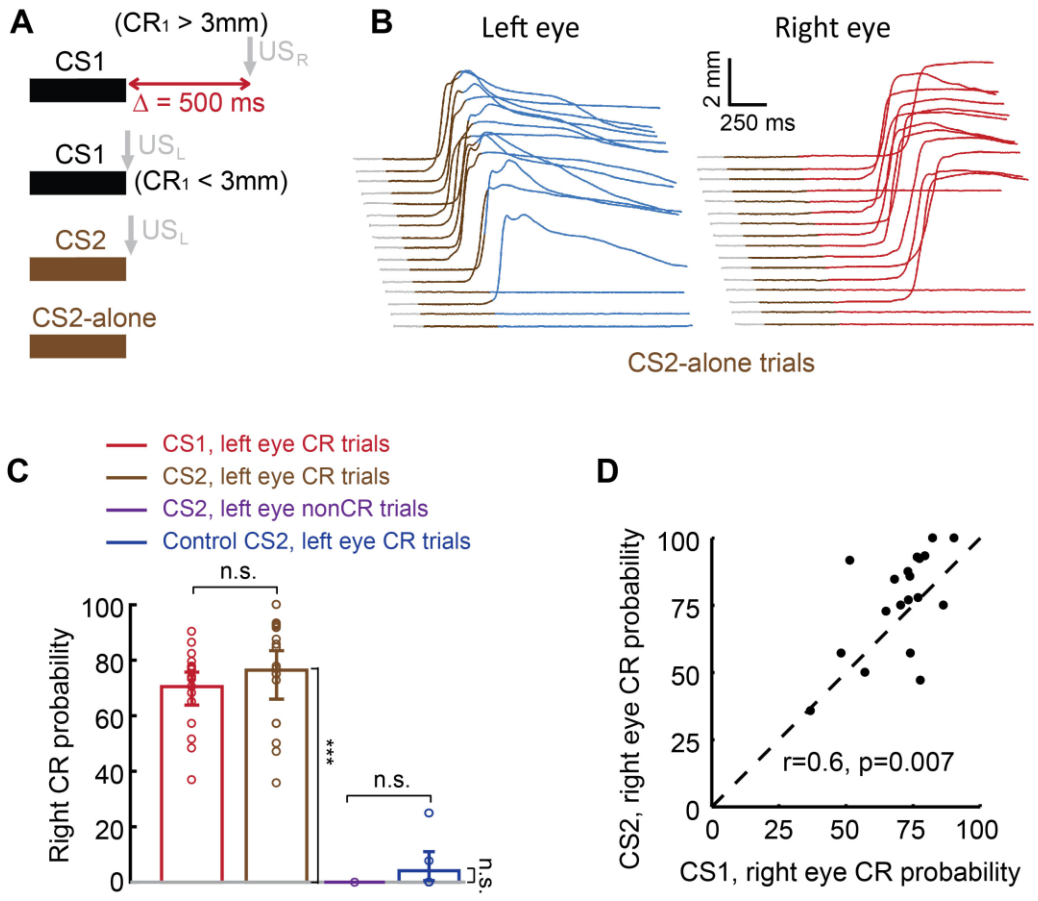

Figure 4.7. Sequence of CRs is present regardless of CS type that drives the first CR (contralateral sequence data).
Same experiment and data presentation as in Figure 4.6 but for subjects trained at contralateral sequence of CRs.

In addition, I repeated CS2-alone test sessions with subjects either not trained with (N=1 and N=3 for ipsi- and contralateral sequence respectively) or extinguished from (N=4 and N=2 for ipsi- and contralateral sequence respectively) producing a sequence of CRs. In these subjects, while first left eye CR amplitudes were larger than 3 mm, CS2 alone trials

did not elicit other CRs in the sequence (ipsilateral sequence: dark blue bars, two-tailed Student t-test, P = 0.33, comparing violet versus dark blue bars with Tukey's post hoc test results in P = 0.99; contralateral sequence: dark blue bars, two-sided Student t-test, P = 0.11, comparing violet versus dark blue bars with Tukey's post hoc test results in P = 0.84). In addition, for subjects trained to produce a sequence of CRs, we observed a significant correlation between probabilities of second CRs in the sequence on CS1 and CS2-alone trials (Fig. 4.6 D and Fig. 4.7 D, r = 0.48, p = 0.026 for ipsilateral sequence, r = 0.6, p = 0.007 for contralateral sequence). Thus, on sessions with better performance of the sequence of CRs on CS1 trials, subjects also showed larger probability of producing a sequence of CRs on CS2-alone trials. Together, these experiments demonstrate that while the first CR in the sequence is driven by the mossy fiber stimulation CS, CRs that follow are driven by FS from the previous CR.

4.4.7 Eyelid Purkinje cells encode similarly all CRs in the sequence

Cerebellar cortex has previously been shown to be necessary for acquisition and expression of well-timed eyelid CRs (Garcia and Mauk, 1998; Kalmbach et al., 2010a). PCs, the sole principal neurons of cerebellar cortex, have been shown to encode the kinematic features of eyelid CRs (Halverson et al., 2015; ten Brinke et al., 2015). Training with two different ISIs within the same session to the same CS also resulted in PCs acquiring two corresponding responses(Halverson et al., 2015; Jirenhed et al., 2017).

I therefore investigated PCs activity during the ipsilateral sequence protocol as a gateway into the underlying neural mechanism. Indeed, if the second CRs are mediated by the usual cerebellar mechanisms, simply driven by a FS rather than CS, then several predictions should follow. First, the same PCs should control the kinematics of both first and subsequent CRs, since both responses are produced by the same muscle. Second, the relationship between PC activity and both responses timing and kinematics should be similar. I tested these predictions with in-vivo recordings and analysis described below.

I chronically implanted in three subjects tetrode micro-drives in the region of cerebellar cortex previously shown to be necessary for acquisition and expression of well-timed eyelid CRs. Subjects were trained in the ipsilateral sequence of CRs using, as before, with electrical stimulation of mossy fibers as the CS. I recorded 156 well-isolated single units during the ipsilateral CRs sequence training sessions. Out of those, 26 were classified as non-eyelid PCs and 16 as eyelid PCs (Halverson et al., 2015) based on US-evoked complex spikes. Data from an example recording session is shown in Fig. 4.8 A, with eyelid CR profiles at the bottom and the eyelid PC raster plot and corresponding PSTH at the top. Even this single example demonstrates that the same PC develops a decrease in activity corresponding in time to each of CRs in the ipsilateral sequence, supporting the first prediction.

To examine whether PC activity drives the expression of the first and second CRs in the ipsilateral sequence, I studied PCs activity aligned to the onset of first and second

CRs (Fig. 4.8 B and C). Figure 4.8 B shows first CR time profile (top) and corresponding PCs activity (bottom) aligned to the first CR onset time (black vertical line). PCs activity was calculated on the first CR trials (blue line) and non-CR trials (black line). Non-CR trials were aligned by CR onset by randomly sampling from the distribution of first CR onset times. Similar to published results during conventional eyelid conditioning sessions (Halverson et al., 2015) and Fig. 3.6 D, PCs activity on CR trials demonstrated a robust decrease in activity prior to CR onset. I next performed a similar analysis using onset times of second CRs, with results shown in Fig. 4.8 C. PC activity on trials with second CRs (blue line) reliably separated from activity on trials without second CRs (black line) prior to the second CR onset. The amount of decrease in PCs activity from baseline at the moment of second CR onset was similar to what was observed for the first CR.

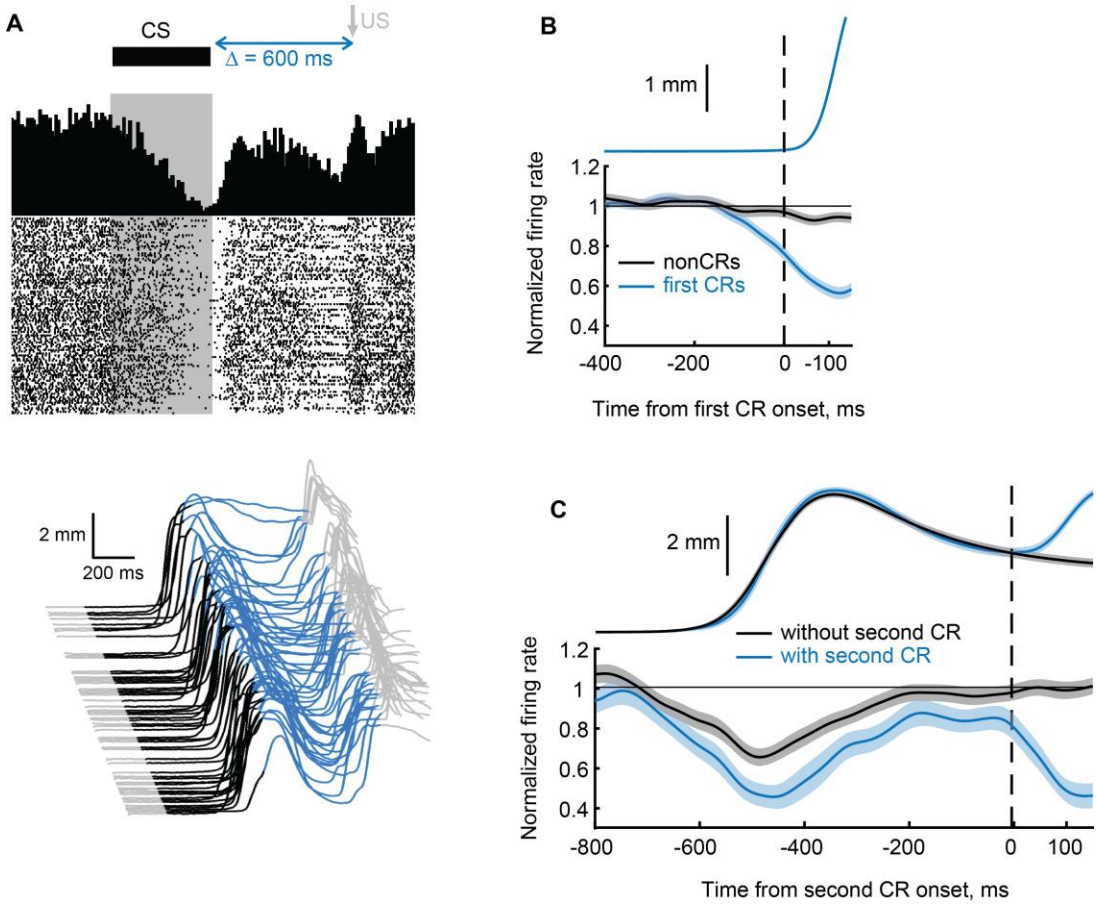

Figure 4.8 Recordings from eyelid Purkinje cells during ipsilateral sequence sessions.
A) Example eyelid PC recording during an ipsilateral sequence training session. Behavioral responses are shown at the bottom, PSTH and a raster plot of eyelid PC are shown at the top. Only trials with first CR amplitude larger than 3 mm are shown in waterfall plot. CS duration is indicated by a grey shaded area. B) Eyelid PCs activity aligned by first CR onset times (vertical black dotted line) is shown at the bottom; aligned CRs profile is shown on top. Eyelid PCs activity on CR trials is shown in blue color, on non-CR trials – in black. C) Behavioral responses and eyelid PCs activity aligned to the onset time of second CR (vertical black dotted line). First CRs are present in all trials, results from trials with second CRs present are shown in blue color, without second CRs – in black.

Next I investigated whether the relationship between eyelid PC activity and CRs timing is the same for both CRs in the sequence. For that I separated trials in three groups based on eyelid CR onset times (non-CRs, early CRs and late CRs) and calculated corresponding average firing rates of eyelid PCs. Results related to the first CR timing are shown in Fig. 4.9 A. Timing of the decrease in eyelid PCs firing rate corresponded to the timing of first CRs, similar and consistent with published results (Halverson et al., 2015; ten Brinke et al., 2015). Namely, on non-CR trials (black lines) eyelid PCs firing rate barely deviated from baseline activity. On trials with early first CRs (red lines) decreases in eyelid PCs firing rate happened earlier than decreases observed on late first CR trials (blue lines). A more important test here was whether the same relationship will hold for the second CRs in the sequence and PCs activity. Results demonstrating that the timing of the PCs decrease also matches the timing of second CRs are shown in Fig. 4.9 B. Here I similarly separated trials into three groups, now based on the presence and timing of the second CRs. Now on trials without the second CRs (black lines), PCs firing rate returned to the baseline level after the decrease corresponding to the first CR. On trials with second CRs present (red and blue lines) PCs activity demonstrated double decreases corresponding to the first and second CRs in the sequence. Moreover, the timing of the second decrease in PCs activity followed the timing of second CRs (Fig. 4.9 B, early versus late second CRs shown by red and blue lines correspondingly). Interestingly, on trials without the second CR the amount of PCs decrease in firing corresponding to the first CR was smaller than on trials with both CRs present.

The results above demonstrate that the same PCs show decreases in activity that precede the onset and match the presence and timing of both eyelid CRs in the ipsilateral sequence. I next examined if this similarity also holds for the relationship between eyelid CR amplitude and the amount of decrease observed in PCs activity. To quantify this relationship, on each trial I calculated PC spike count during the last 300 ms prior to the US (US_1 time for analysis of first CRs, US_2 time for analysis of second CRs) and normalized it to the average spike count during baseline activity. Values around one correspond to no change from baseline firing, values below one indicate a decrease in PC activity and values above one indicate an increase in firing relative to the baseline level. Figure 4.9 C shows PCs spike count as a function of the first CR amplitude, with each dot corresponding to a single trial. I found a robust correlation between the amount of decrease in PCs activity and first CR amplitude (Pearson correlation coefficient $r = -0.27$, $P = 1 \cdot 10^{-20}$). Similar analysis done for the second CR trials is shown in Fig. 4.9 D. Again, I found a robust correlation between the amount of decrease in PCs firing and second CR amplitude ($r = -0.42$, $P = 3 \cdot 10^{-24}$). Overall, recordings from eyelid PCs demonstrate that the same cerebellar mechanisms are involved in generation of both CRs in the sequence. Consistent with these data is the general notion that FS can be used by the cerebellum to learn following responses in the sequence through processes similar to the learning of the first response using a CS.

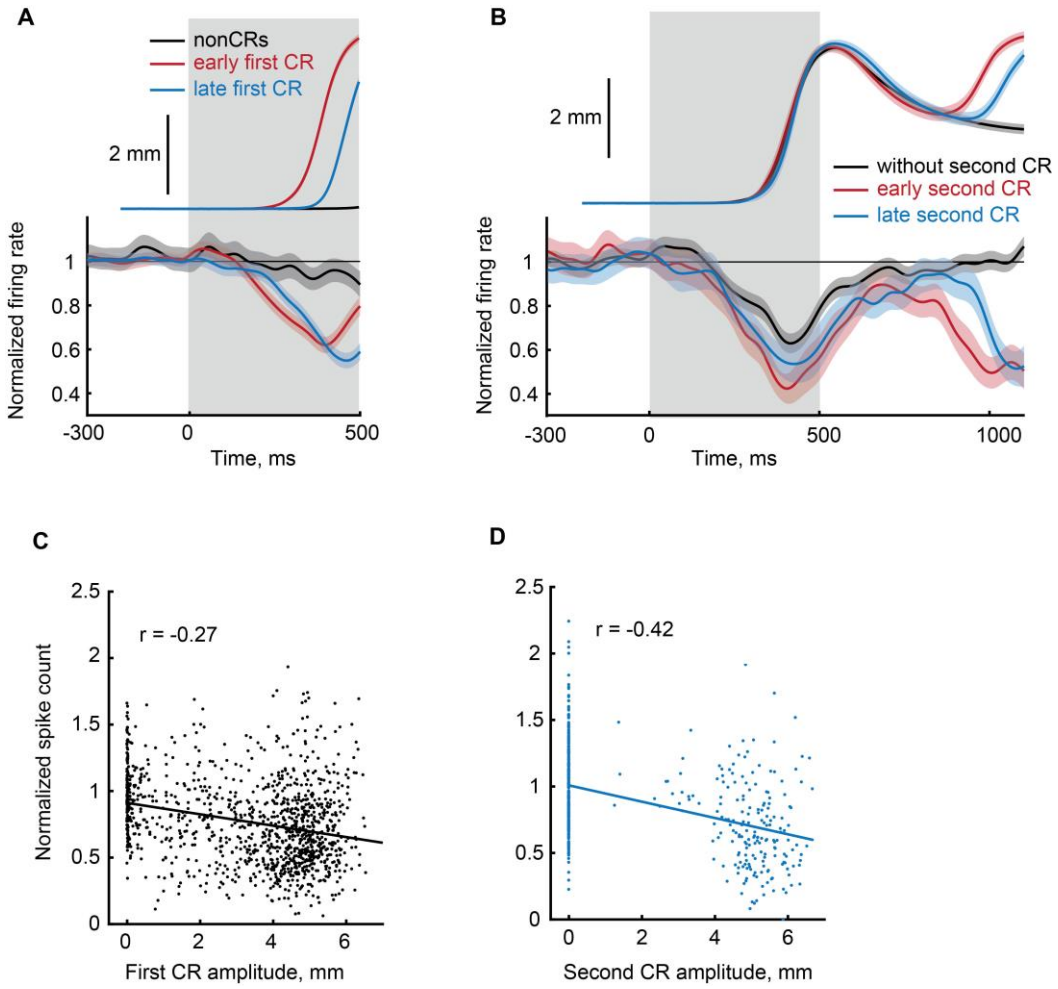

Figure 4.9. Eyelid PCs control the timing and amplitude of both responses in the sequence.
A) Average eyelid response profiles on trials sorted by the onset time of first CR are shown on top, corresponding average eyelid PCs firing rate normalized by the baseline level is shown at the bottom. Non-CR trials are shown in black, trials with early and late CR onset times are shown in red and blue respectively. Behavioral responses and eyelid PCs activity are truncated at US_1 onset, shaded regions represent 95% confidence intervals. B) Similar to A), but for trials with first CR present and sorted by the onset times of second CR. Color-coding is similar to A), behavioral responses and eyelid PCs activity are truncated at US_2 onset.
C) Spike count of eyelid PCs activity during last 300 ms before US_1 time versus first CR amplitude. Spike counts are normalized to the average spike count during baseline activity. Each dot represents a single trial; a solid black line represents the best linear fit. D) Normalized spike count of eyelid PCs activity during last 300 ms before US_2 time versus second CR amplitude. Each dot represents a single trial; a solid blue line represents the best linear fit.

4.4 Discussion

In this chapter I investigated how movement sequences are learned and implemented by the cerebellum. I demonstrated that the feedback information from the first CR is a sufficient signal for the cerebellum to learn the next response in the sequence. Results held true for both ipsilateral and contralateral sequence of eyelid CRs. Through a series of experiments I showed that responses following the first CR were associated with a feedback signal from it and not with the sensory input to the cerebellum provided by electrical stimulation of mossy fiber in our experiments. Finally, recordings from eyelid PCs, the sole output neurons of cerebellar cortex, show that all CRs in the sequence are encoded in the same manner by the cerebellum. In summary, these results demonstrate how, through a simple associative learning procedure, the cerebellum can learn to chain together a desired sequence of appropriately timed movements.

While I have demonstrated the sufficiency of FS from a CR to be used by the cerebellum as a new "CS", the pathway of such FS remains an open question. Several pathway routes are possible. The most direct feedback route would originate from the cerebellar output region — the deep cerebellar nucleus neurons (DCN). DCN neurons are known to increase their activity during CR expression (McCormick and Thompson, 1984a; Halverson et al., 2010). Information about a CR can be passed back to granule cells in the cerebellar cortex via a monosynaptic excitation by axon collaterals of glutamatergic DCN neurons (Houck and Person, 2015; Gao et al., 2016) or via a

reduction of inhibition from Golgi cells that receive inhibitory projections (Ankri et al., 2015) from GABA/glycinergic DCN neurons. More indirect routes to convey DCN activity back to cerebellar cortex also exist, an example being via thalamic and pontine nuclei neurons (Halverson et al., 2010). The FS could also originate not directly from the cerebellar output activity, but rather from proprioceptive information driven by the movement itself. All of these possible routes can contribute to a FS and their relative contributions will likely depend on temporal constrains and/or laterality of the movement sequence. For the parameters used in the ipsilateral sequence protocol described above, Fig. 4.9 B provides evidence that the direct cerebellar output constitutes a large portion of the FS. In this scenario, a smaller decrease in PCs firing during the first CR will lead to a smaller disinhibition of DCN neurons and consequently to a 'weaker' FS received by the cerebellum. This, according to results described in Chapter 3, should in turn lead to a smaller probability of observing the second CR, which is exactly was I observed. In general however, while we know that the feedback information is supplied to the cerebellar cortex (Giovannucci et al., 2017), further studies are needed to investigate the relative contributions of possible feedback pathways.

In these experiments I trained subjects to produce a sequence of discreet responses. The results of this study however should not rely on the discreetness of responses, which I utilized for the clearness of results. Natural complex multi-joint movements can be broken down into several components, though such separations could be to a degree arbitrary since components of the movement smoothly transition into each

other. In the case of natural movements our results simply imply that the feedback about the initial portion of the movement can be used by the cerebellum to learn a correct continuation of the movement sequence.

The cerebellum and cerebral cortex areas, such as premotor and primary motor cortices, are known to form a closed loop system via cortico-ponto-cerebellar projections(Evarts and Thach, 1969; Glickstein et al., 1985; Kelly and Strick, 2003), where the cerebellum sends information back to the cerebral cortex via the thalamus. Electrical (Penfield and Boldrey, 1937; Graziano et al., 2002) or optogenetic (Harrison et al., 2012) stimulation of motor cortex area is known to produce complex multi-joint movements. A common interpretation of these results is a notion of motor program (Mink, 1996; Summers and Anson, 2009), stored either entirely within a motor cortex or at least partially within downstream areas. Our study provides new evidence towards interpretation that motor program could be stored in part within the cerebellum. While the initial command from motor cortex initiates the movement, the later feedback signal from cerebellar output and the movement itself is used by the cerebellum to learn a proper output that modifies the activity in the motor cortex so that the movement trajectory is close to the desired target trajectory. While speculative at this point, such a framework is fully compatible with results of the present study.

CHAPTER 5

Final remarks

From the engineering perspective, in order to study the computation of a particular brain system, one would want the ability to control the inputs and measure the output. In neuroscience, such an approach could be considered naïve and often simply not feasible due to massive interconnectivity between and within different bran regions. The cerebellum might be one of the unique places in the brain, except early sensory systems, where both the control over inputs and behavioral readout of the output are possible. All sensory modalities on inputs are conveyed to the cerebellum via mossy fibers, originating at the pontine nuclei. Luckily for me and other cerebellar researchers, mossy fibers are anatomically constrained to the bundle of white matter called middle cerebellar peduncle. Therefore, a simple technique such as electrical stimulation through electrode implanted in the middle cerebellar peduncle provides the ability to control and constrain inputs to the cerebellum. Eyelid CRs, in turn, serve as a direct quantifiable metric of cerebellar output, as demonstrated in Chapter 2. In all studies reported here I utilized this direct mapping of eyelid conditioning protocol onto cerebellar circuitry to study different aspects of cerebellar computation. I believe that such advantageous setup will be a foundation for many exciting findings in the future.

The results of Chapter 2 show that activity of eyelid PCs is sufficient to predict the time profiles of eyelid CRs. Since PCs are the sole output of cerebellar cortex, that implies that this information is already present at that stage of processing. However, the final output of the cerebellum, DCN neurons, receive apart from PCs inhibitory input also a directly excitatory input from mossy fibers. My results imply that, at least in eyelid conditioning, the direct input does not provide any critical information to DCN neurons that determine the profile of a CR. Meanwhile, the existence of LTP and LTD at the mossy fibers to DCN synapses has been shown by several in-vitro studies (Pugh and Raman, 2006, 2008). Behavioral studies using a pharmacological disassociation of cerebellar cortex inhibition demonstrated a potentiation specific to CS activated mossy fibers to DCN synapses following acquisition of CRs (Medina et al., 2001; Ohyama et al., 2003), consistent with the induction of LTP on those synapses. Moreover, recent studies using optogenetic manipulation of PCs activity to mimic in-vitro LTD and LTD protocols found strong behavioral effects in eyelid-conditioning-like task involving paw movement (Lee et al., 2015) and VOR (Shin et al., 2014). These results can be easily reconciled with mine if the direct mossy fiber input representing CS is to large degree temporally uniform. In this case the learned potentiation of mossy-fiber to DCN synapses would not contribute any temporal information to DCN neurons.

This question can also be addressed directly in future by applying the analysis similar to what I performed on eyelid PCs now on recordings from DCN neurons. A hard part here, apart from obtaining the recordings data, would be to find a way to isolate

"eyelid" DCN neurons. One can potentially use antidromic stimulation to isolate DCN neurons projecting downstream to red nucleus. In order to further constrain neurons to those that drive eyelid CR one can use microsimulation through tetrodes in order to elicit an eyelid movement ref. A conjunction of these two approaches should provide a reasonable exclusion criterion of DCN neurons that would show changes in firing during CS, but are either GABAergic neurons projecting to the inferior olive or are related to non-eyelid movements that may accompany an eyelid CR. The differences in results between predictions of eyelid CRs profiles from eyelid DCN neurons versus eyelid PCs will provide insights about the contribution of a direct mossy fibers input to DCN neurons.

Experiments described in Chapter 3 were designed to mimic the learning of a single correct motor response for a specific input. In this scenario one would indeed expect the responses to a new input to be bimodal: with one mode corresponding to cases where new input is classified as the trained input and with the second mode being non-responses, since we assumed that the cerebellum had learned to produce the responses only to the trained input. Such framework is obviously a simplification of the real world picture. The next step would be to increase "the dimensionality" of input-output space and study does the phenomenon still hold. One could implement that by training subjects to produce different CR amplitudes (e.g. 3 mm and 6 mm) in response to two distinct inputs. After that the test would be to present the probe inputs representing a mixture of two trained inputs at different proportions, similar to the rules of competing stimulus

protocol. A prediction from the current binary choice framework would be to observe three modes of responses: non-responses, 3 mm responses and 6 mm responses, with probabilities corresponding to the fractions of trained inputs from which the probe consists.

There are several possible ways how the cerebellum can implement the binary choice phenomenon on a circuit level. The hypothesis I described in Chapter 3 essentially breaks the implementation into two potentially fully separate mechanisms. The first mechanism governs the initiation of behavioral response, while the second ensures that the final amplitude of the initiated response is correct. In this framework, first granule-Golgi cells network converges to a state that initiates a behavioral response. At the level of eyelid PCs, this would correspond to an initial decrease to about 75 percent of the baseline activity, reported previously (see (Halverson et al., 2015) and Chapter 4 Fig. 4.8 B, C). The disinhibition of DCN neurons results in the increase of their firing rate and initiation of the behavioral response. This in turn initiates a feedback signal to the cerebellar cortex (the second mechanism), either directly thought the DCN neurons axon collaterals or though more indirect routes.

A partially alternative implementation would be if the granule-Golgi cells network acts as an attractor network. This is similar to the proposed first mechanism above, but instead of just converging to the state that initiates a response, here it would need to act as a line attractor (Machens and Brody, 2008; Laje and Buonomano, 2013;

Miller, 2016) to stably converge to the trajectory in state-space that also drives the expression of the response, resulting in a proper amplitude. While none of the experimental data so far contradict this possibility, it is a harder requirement on the network comparing to just initiating the response. The two proposed mechanisms do not have to be fully separate: it can also be that the feedback signal from sub-behavioral threshold increases in DCN firing also participates in the 'decision' of whether there should be a CR or nonCR on a given trial.

All of these possible options can be at least partially disentangled with future experiments. The most direct way would be to ontogenetically silence the axon terminals of DCN axonal projections in cerebellar cortex. Though necessary, such an endeavor will have to overcome several challenges.

First, while optogenetic inhibition of neurons somas is quite robust with modern silencers, the robustness of silencing axon projections has been questioned (Mahn et al., 2016). Even if inhibition of axonal terminals is not a problem, the amount of viral expression and placement of optic fiber can result in silencing only a portion of DCN axon collaterals involved in the task. To partially overcome this challenge one can include a large number of subjects in the experiment and correlate the size of the behavioral effect with the amount of viral expression and optic fiber placement.

Second, the parameters of the behavioral task can play a major role on the amount of DCN feedback contribution to CR expression. I hypothesized that DCN feedback is

used by the cerebellum to learn CR amplitude, as opposed to being a fixed positive feedback. Thus the amount of training subjects receive will influence the amount of DCN feedback contribution. One would expect to observe large effects only in subjects that had acquired robust CRs with large, stable amplitudes. In addition, training parameters like ISI duration can have large influence. Results from parallel fiber LTD protocols state that granule cells need to be active about 100 ms prior to climbing fiber input for a successful LTD induction. If ISI is short (about 200-250 ms), then granule cell activated by the feedback will fire too late for LTD induction. Obviously, the numbers used in this arguments are simplification of the underlying complex processes. Nevertheless, one would expect to observe a diminishing effect of DCN feedback with the decrease in ISI used for training. Consistent with that, one can make a prediction that 'the binarity' of the binary choice should break if subjects will be trained and tested at short ISIs. Our pilot data (not shown) confirms this prediction, but I would hope to see more direct tests of DCN feedback contribution to binary choice phenomenon in the future.

While the phenomena addressed in Chapters 3 and 4 are behaviorally quite different, there is interestingly a unifying aspect of cerebellar physiology that they might rely on. Results of computer simulations described in Chapter 3 suggest that feedback from DCN neurons is crucial for implementation of binary response amplitudes (in target-or-none sense). Similarly, in Chapter 4 I demonstrated that the feedback information about a CR is sufficient for the cerebellum to learn the next response in the

sequence. The hypothesis stated in Chapter 3 is quite similar, with the difference being that the feedback is used within the response rather than in the next movement. These results highlight that effect of feedback signals should not be overlooked even in such seemingly feed-forward network as the cerebellum.

References

A. Russo, B. London, S. Perkins MC (2016) Changes in motor cortex population structure between movement types. In: Society for Neuroscience Conference. San Diego, CA.

Albus JS (1971) A theory of cerebellar function. Math Biosci 10:25–61.

Ankri L, Husson Z, Pietrajtis K, Proville R, Léna C, Yarom Y, Dieudonné S, Uusisaari MY (2015) A novel inhibitory nucleo-cortical circuit controls cerebellar Golgi cell activity. Elife 4:1–26.

Asanuma C, Thach WT, Jones EG (1983a) Distribution of cerebellar terminations and their relation to other afferent terminations in the ventral lateral thalamic region of the monkey. Brain Res 286:237–265.

Asanuma C, Thach WT, Jones EG (1983b) Brainstem and spinal projections of the deep cerebellar nuclei in the monkey, with observations on the brainstem projections of the dorsal column nuclei. Brain Res 286:299–322.

Bastian AJ (2006) Learning to predict the future: the cerebellum adapts feedforward movement control. Curr Opin Neurobiol 16:645–649.

Best AR, Regehr WG (2009) Inhibitory Regulation of Electrically Coupled Neurons in the Inferior Olive Is Mediated by Asynchronous Release of GABA. Neuron 62:555–565.

Billings G, Piasini E, Lőrincz A, Nusser Z, Silver RA (2014) Network Structure within the Cerebellar Input Layer Enables Lossless Sparse Encoding. Neuron 83:960–974.

Boyd J, Aitkin L (1976) Responses of single units in the pontine nuclei of the cat to

acoustic stimulation. Neurosci Lett 3:259–263.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (n.d.) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 13:87–100.

Brodal P (1968) The corticopontine projection in the cat. I. Demonstration of a somatotopically organized projection from the primary sensorimotor cortex. Exp brain Res 5:210–234.

Buonomano D V., Mauk MD (1994) Neural Network Model of the Cerebellum: Temporal Discrimination and the Timing of Motor Responses. Neural Comput 6:38–55.

Campolattaro MM, Halverson HE, Freeman JH (2007) Medial auditory thalamic stimulation as a conditioned stimulus for eyeblink conditioning in rats. Learn Mem 14:152–159.

Carrillo J, Nishiyama N, Nishiyama H (2013) Dendritic Translocation Establishes the Winner in Cerebellar Climbing Fiber Synapse Elimination. J Neurosci 33:7641–7653.

Castiello U, Bennett KM, Stelmach GE (1993) The bilateral reach to grasp movement. Behav Brain Res 56:43–57.

Chapman PF, Steinmetz JE, Sears LL, Thompson RF (1990) Effects of lidocaine injection in the interpositus nucleus and red nucleus on conditioned behavioral and neuronal responses. Brain Res 537:149–156.

Cicirata F, Serapide MF, Parenti R, Pant? MR, Zappal? A, Nicotra A, Cicero D (2005) The basilar pontine nuclei and the nucleus reticularis tegmenti pontis subserve

distinct cerebrocerebellar pathways. In: Progress in brain research, pp 259–282.

Desclin JC (1974) Histological evidence supporting the inferior olive as the major source of cerebellar climbing fibers in the rat. Brain Res 77:365–384.

Desmurget M, Turner RS (2010) Motor sequences and the basal ganglia: kinematics, not habits. J Neurosci 30:7685–7690.

Diener H-C, Dichgans J (1992) Pathophysiology of cerebellar ataxia. Mov Disord 7:95–109.

Doyon J, Gaudreau D, Jr. RL, Castonguay M, Bédard PJ, Bédard F, Bouchard J-P (1997) Role of the Striatum, Cerebellum, and Frontal Lobes in the Learning of a Visuomotor Sequence. Brain Cogn 34:218–245.

Doyon J, Song AW, Karni A, Lalonde F, Adams MM, Ungerleider LG (2002) Experience-dependent changes in cerebellar contributions to motor sequence learning. Proc Natl Acad Sci 99:1017–1022.

Eccles JC (1967) Circuits in the cerebellar control of movement. Proc Natl Acad Sci U S A 58:336–343.

Eilers J, Callewaert G, Armstrong C, Konnerth A (1995) Calcium signaling in a narrow somatic submembrane shell during synaptic activity in cerebellar Purkinje neurons. Proc Natl Acad Sci U S A 92:10272–10276.

Evarts E V, Thach WT (1969) Motor Mechanisms of the CNS: Cerebrocerebellar Interrelations. Annu Rev Physiol 31:451–498.

Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc Am A 4:2379–2394.

Flumerfelt BA, Otabe S, Courville J (1973) Distinct projections to the red nucleus from the dentate and interposed nuclei in the monkey. Brain Res 50:408–414.

Freeman JH, Nicholson DA, Muckler AS, Rabinak CA, DiPietro NT (2003) Ontogeny of eyeblink conditioned response timing in rats. Behav Neurosci 117:283–291.

Gall D (2005) Intracellular Calcium Regulation by Burst Discharge Determines Bidirectional Long-Term Synaptic Plasticity at the Cerebellum Input Stage. J Neurosci 25:4813–4822.

Gao H, Solages C de, Lena C (2012) Tetrode recordings in the cerebellar cortex. J Physiol Paris 106:128–136.

Gao Z, Proietti-Onori M, Lin Z, ten Brinke MM, Boele HJ, Potters JW, Ruigrok TJH, Hoebeek FE, De Zeeuw CI (2016) Excitatory Cerebellar Nucleocortical Circuit Provides Internal Amplification during Associative Conditioning. Neuron 89:645–657.

Garcia KS, Mauk MD (1998) Pharmacological analysis of cerebellar contributions to the timing and expression of conditioned eyelid responses. Neuropharmacology 37:471–480.

Garcia KS, Steele PM, Mauk MD (1999) Cerebellar cortex lesions prevent acquisition of conditioned eyelid responses. J Neurosci 19:10940–10947.

Giovannucci A, Badura A, Deverett B, Najafi F, Pereira TD, Gao Z, Ozden I, Kloth AD, Pnevmatikakis E, Paninski L, De Zeeuw CI, Medina JF, Wang SS-H (2017) Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. Nat Neurosci 20:727–734.

Glickstein M, Cohen JL, Dixon B, Gibson A, Hollins M, Labossiere E, Robinson F

(1980) Corticopontine visual projections in macaque monkeys. J Comp Neurol 190:209–229.

Glickstein M, May JG, Mercier BE (1985) Corticopontine projection in the macaque: The distribution of labelled cortical cells after large injections of horseradish peroxidase in the pontine nuclei. J Comp Neurol 235:343–359.

Grafton ST, Hazeltine E, Ivry RB (2002) Motor sequence learning with the nondominant left hand: A PET functional imaging study. Exp Brain Res 146:369–378.

Graziano MSA, Taylor CSR, Moore T, Sencer W, Hamuy TP, Travis AM (2002) Complex movements evoked by microstimulation of precentral cortex. Neuron 34:841–851.

Guo CC, Ke MC, Raymond JL (2014) Cerebellar Encoding of Multiple Candidate Error Cues in the Service of Motor Learning. J Neurosci 34:9880–9890.

Halverson HE, Freeman JH (2010) Medial auditory thalamic input to the lateral pontine nuclei is necessary for auditory eyeblink conditioning. Neurobiol Learn Mem 93:92–98.

Halverson HE, Hubbard EM, Freeman JH (2009) Stimulation of the lateral geniculate, superior colliculus, or visual cortex is sufficient for eyeblink conditioning in rats. Learn Mem 16:300–307.

Halverson HE, Khilkevich A, Mauk MD (2015) Relating cerebellar purkinje cell activity to the timing and amplitude of conditioned eyelid responses. J Neurosci 35:7813–7832.

Halverson HE, Lee I, Freeman JH (2010) Associative plasticity in the medial auditory thalamus and cerebellar interpositus nucleus during eyeblink conditioning. J

Neurosci 30:8787–8796.

Hansel C, de Jeu M, Belmeguenai A, Houtman SH, Buitendijk GS, Andreev D, De?Zeeuw C, Elgersma Y (2006) ?CaMKII Is Essential for Cerebellar LTD and Motor Learning. Neuron 51:835–843.

Hansel C, Linden DJ (2000) Long-term depression of the cerebellar climbing fiber--Purkinje neuron synapse. Neuron 26:473–482.

Harrison TC, Ayling OGS, Murphy TH, Liewald JF, Kay K, Watzke N, Wood PG, Bamberg E, Nagel G, Gottschalk A, Deisseroth K (2012) Distinct cortical circuit mechanisms for complex forelimb movement and motor map topography. Neuron 74:397–409.

Hashimoto K, Ichikawa R, Kitamura K, Watanabe M, Kano M (2009) Translocation of a ?Winner? Climbing Fiber to the Purkinje Cell Dendrite and Subsequent Elimination of ?Losers? from the Soma in Developing Cerebellum. Neuron 63:106–118.

Hazeltine E, Ivry RB (2002) Neuroscience. Can we teach the cerebellum new tricks? Science 296:1979–1980.

Heiney SA, Kim J, Augustine GJ, Medina JF (2014) Precise Control of Movement Kinematics by Optogenetic Inhibition of Purkinje Cell Activity. J Neurosci 34:2321–2330.

Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2015) Encoding of action by the Purkinje cells of the cerebellum. Nature 526:439–442.

Hesslow G, Ivarsson M (1994) Suppression of cerebellar Purkinje cells during conditioned responses in ferrets. Neuroreport 5:649–652.

Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442:164–171.

Houck BD, Person AL (2014) Cerebellar loops: a review of the nucleocortical pathway. Cerebellum 13:378–385.

Houck BD, Person AL (2015) Cerebellar premotor output neurons collateralize to innervate the cerebellar cortex. J Comp Neurol 523:2254–2271.

Hull C, Regehr WG (2012) Identification of an inhibitory circuit that regulates cerebellar Golgi cell activity. Neuron 73:149–158.

Ito M (1984) The modifiable neuronal network of the cerebellum. Jpn J Physiol 34:781–792.

Jackson GM, Jackson SR, Harrison J, Henderson L, Kennard C (1995) Serial reaction time learning and Parkinson's disease: evidence for a procedural learning deficit. Neuropsychologia 33:577–593.

Jirenhed D-A, Hesslow G (2011) Time course of classically conditioned Purkinje cell response is determined by initial part of conditioned stimulus. J Neurosci 31:9070–9074.

Jirenhed D-A, Rasmussen A, Johansson F, Hesslow G (2017) Learned response sequences in cerebellar Purkinje cells. Proc Natl Acad Sci U S A 114:6127–6132.

Jirenhed D-AD-A, Bengtsson F, Hesslow G (2007) Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. J Neurosci 27:2493–2502.

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999) Building neural representations of habits. Science 286:1745–1749.

Kalmbach BE, Davis T, Ohyama T, Riusech F, Nores WL, Mauk MD (2010a) Cerebellar Cortex Contributions to the Expression and Timing of Conditioned Eyelid Responses. J Neurophysiol 103:2039–2049.

Kalmbach BE, Ohyama T, Kreider JC, Riusech F, Mauk MD (2009) Interactions between prefrontal cortex and cerebellum revealed by trace eyelid conditioning. Learn Mem 16:86–95.

Kalmbach BE, Ohyama T, Mauk MD (2010b) Temporal Patterns of Inputs to Cerebellum Necessary and Sufficient for Trace Eyelid Conditioning. J Neurophysiol 104:627–640.

Kalmbach BE, Voicu H, Ohyama T, Mauk MD (2011) A Subtraction Mechanism of Temporal Coding in Cerebellar Cortex. J Neurosci 31:2025–2034.

Kano M, Hashimoto K (2009) Synapse elimination in the central nervous system. Curr Opin Neurobiol 19:154–161.

Keating JG, Thach WT (1995) Nonclock behavior of inferior olive neurons: interspike interval of Purkinje cell complex spike discharge in the awake behaving monkey is random. J Neurophysiol 73:1329–1340.

Kelly RM, Strick PL (2003) Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J Neurosci 23:8432–8444.

Kelso J, Southard D, Goodman D (1979) On the nature of human interlimb coordination. Science (80-) 203.

Kenyon GT (1997) A model of long-term memory storage in the cerebellar cortex: a possible role for plasticity at parallel fiber synapses onto stellate/basket interneurons. Proc Natl Acad Sci U S A 94:14200–14205.

Kenyon GT, Medina JF, Mauk MD (1998) A mathematical model of the cerebello-olivary system. I: Self-regulating equilibrium of climbing fiber activity. J Comput Neurosci 5:17–33.

Khilkevich A, Halverson HE, Canton-Josh JE, Mauk MD (2016) Links Between Single-Trial Changes and Learning Rate in Eyelid Conditioning. Cerebellum 15:112–121.

Kim J, Lee S, Tsuda S, Zhang X, Asrican B, Gloss B, Feng G, Augustine GJ (2014) Optogenetic Mapping of Cerebellar Inhibitory Circuitry Reveals Spatially Biased Coordination of Interneurons via Electrical Synapses. Cell Rep 7:1601–1613.

Kitama T, Omata T, Mizukoshi A, Ueno T, Sato Y (1999) Motor Dynamics Encoding in Cat Cerebellar Flocculus Middle Zone During Optokinetic Eye Movements. J Neurophysiol 82.

Kratochwil CF, Maheshwari U, Rijli FM (2017) The Long Journey of Pontine Nuclei Neurons: From Rhombic Lip to Cortico-Ponto-Cerebellar Circuitry. Front Neural Circuits 11:33.

Kreider JC, Mauk MD (2010) Eyelid conditioning to a target amplitude: adding how much to whether and when. J Neurosci 30:14145–14152.

Krupa DJ, Thompson JK, Thompson RF (1993) Localization of a memory trace in the mammalian brain. Science (80-) 260:989–991.

Krupa DJ, Thompson RF (1995) Inactivation of the superior cerebellar peduncle blocks expression but not acquisition of the rabbit's classically conditioned eye-blink response. Proc Natl Acad Sci U S A 92:5097–5101.

Krupa DJ, Thompson RF (1997) Reversible inactivation of the cerebellar interpositus nucleus completely prevents acquisition of the classically conditioned eye-blink

response. Learn Mem 3:545–556.

Krupa DJ, Weng J, Thompson RF (1996) Inactivation of brainstem motor nuclei blocks expression but not acquisition of the rabbit's classically conditioned eyeblink response. Behav Neurosci 110:219–227.

Laje R, Buonomano D V (2013) Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat Neurosci 16:925–933.

Lang EJ, Sugihara I, Welsh JP, Llinás R (1999) Patterns of spontaneous purkinje cell complex spike activity in the awake rat. J Neurosci 19:2728–2739.

Lee K, Mathews P, Reeves AB, Choe K, Jami S, Serrano R, Otis T (2015) Circuit Mechanisms Underlying Motor Memory Formation in the Cerebellum. Neuron 86:529–540.

Leergaard TB, Lillehaug S, De Schutter E, Bower JM, Bjaalie JG (2006) Topographical organization of pathways from somatosensory cortex through the pontine nuclei to tactile regions of the rat cerebellar hemispheres. Eur J Neurosci 24:2801–2812.

Lefler Y, Yarom Y, Uusisaari MY (2014) Cerebellar Inhibitory Input to the Inferior Olive Decreases Electrical Coupling and Blocks Subthreshold Oscillations. Neuron 81:1389–1400.

Lehéricy S, Benali H, Van de Moortele P-F, Pélégrini-Issac M, Waechter T, Ugurbil K, Doyon J (2005) Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. Proc Natl Acad Sci U S A 102:12566–12571.

Lepora NF, Mavritsaki E, Porrill J, Yeo CH, Evinger C, Dean P (2007) Evidence From Retractor Bulbi EMG for Linearized Motor Control of Conditioned Nictitating Membrane Responses. J Neuro-physiol 98:2074–2088.

Lepora NF, Porrill J, Yeo CH, Evinger C, Dean P (2009) Recruitment in Retractor Bulbi Muscle During Eyeblink Conditioning: EMG Analysis and Common-Drive Model. J Neurophysiol 102:2498–2513.

Li W-K, Hausknecht MJ, Stone P, Mauk MD (2013) Using a million cell simulation of the cerebellum: Network scaling and task generality. Neural Networks 47:95–102.

Lisberger SG (1994) Neural Basis for Motor Learning in the Vestibuloocular Reflex of Primates .3. Computational and Behavioral-Analysis of the Sites of Learning. J Neurophysiol 72:974–998.

Liu S, Gu Y, DeAngelis GC, Angelaki DE (2012) Choice-related activity and correlated noise in subcortical vestibular neurons. Nat Neurosci 16:89–97.

Machens CK, Brody CD (2008) Design of Continuous Attractor Networks with Monotonic Tuning Using a Symmetry Principle. Neural Comput 20:452–485.

Maekawa K, Kimura M, Takeda T (1981) Mossy fiber activation of the cerebellar flocculus from the visual system. Ann N Y Acad Sci 374:476–490.

Mahn M, Prigge M, Ron S, Levy R, Yizhar O (2016) Biophysical constraints of optogenetic inhibition at presynaptic terminals. Nat Neurosci 19:554–556.

Marr D (1969) A theory of cerebellar cortex. J Physiol 202:437–470.

Mauk MD, Donegan NH (1997) A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. Learn Mem 4:130–158.

Mauk MD, Ohyama T (2004) Extinction as New Learning Versus Unlearning: Considerations from a Computer Simulation of the Cerebellum. Learn Mem 11:566–571.

Mauk MD, Ruiz BP (1992) Learning-dependent timing of Pavlovian eyelid responses: differential conditioning using multiple interstimulus intervals. Behav Neurosci 106:666–681.

Mauk MD, Steinmetz JE, Thompson RF (1986) Classical conditioning using stimulation of the inferior olive as the unconditioned stimulus. Proc Natl Acad Sci U S A 83:5349–5353.

McCormick DA, Steinmetz JE, Thompson RF (1985) Lesions of the inferior olivary complex cause extinction of the classically conditioned eyeblink response. Brain Res 359:120–130.

McCormick DA, Thompson RF (1984a) Neuronal responses of the rabbit cerebellum during acquisition and performance of a classically conditioned nictitating membrane-eyelid response. J Neurosci 4:2811–2822.

McCormick DA, Thompson RF (1984b) Cerebellum: essential involvement in the classically conditioned eyelid response. Science 223:296–299.

Medina JF, Garcia KS, Mauk MD (2001) A mechanism for savings in the cerebellum. J Neurosci 21:4081–4089.

Medina JF, Garcia KS, Nores WL, Taylor NM, Mauk MD (2000) Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. J Neurosci 20:5516–5525.

Medina JF, Lisberger SG (2007) Variation, signal, and noise in cerebellar sensory-motor processing for smooth-pursuit eye movements. J Neurosci Off J Soc Neurosci 27:6832–6842.

Medina JF, Lisberger SG (2009) Encoding and decoding of learned smooth-pursuit eye

movements in the floccular complex of the monkey cerebellum. J Neurophysiol 102:2039–2054.

Medina JF, Mauk MD (1999) Simulations of cerebellar motor learning: Computational analysis of plasticity at the mossy fiber to deep nucleus synapse. Struct Large-Scale Synchronized Firing Primate Retin 19:7140–7151.

Medina JF, Nores WL, Mauk MD (2002) Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. Nature 416:330–333.

Millenson JR, Kehoe EJ, Gormezano I (1977) Classical conditioning of the rabbit's nictitating membrane response under fixed and mixed CS-US intervals. Learn Motiv 8:351–366.

Miller P (2016) Dynamical systems, attractors, and neural circuits. F1000Research 5:992.

Mink JW (1996) THE BASAL GANGLIA: FOCUSED SELECTION AND INHIBITION OF COMPETING MOTOR PROGRAMS. Prog Neurobiol 50:381–425.

O'Reilly RC, McClelland JL (1994) Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. Hippocampus 4:661–682.

Ohmae S, Medina JF (2015) Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. Nat Neurosci 29:1–7.

Ohyama T, Mauk M (2001) Latent acquisition of timed responses in cerebellar cortex. J Neurosci 21:682–690.

Ohyama T, Nores WL, Mauk MD (2003) Stimulus generalization of conditioned eyelid responses produced without cerebellar cortex: implications for plasticity in the cerebellar nuclei. Learn Mem 10:346–354.

Ohyama T, Nores WL, Medina JF, Riusech FA, Mauk MD (2006) Learning-Induced Plasticity in Deep Cerebellar Nucleus. J Neurosci 26:12656–12663.

Ohyama T, Voicu H, Kalmbach B, Mauk MD (2010) A Decrementing Form of Plasticity Apparent in Cerebellar Learning. J Neurosci 30:16993–17003.

Omata T, Kitama T, Mizukoshi A, Ueno T, Kawato M, Sato Y (2000) Purkinje cell activity in the middle zone of the cerebellar flocculus during optokinetic and vestibular eye movement in cats. Jpn J Physiol 50:357–370.

Otmakhov N, Griffith LC, Lisman JE (1997) Postsynaptic inhibitors of calcium/calmodulin-dependent protein kinase type II block induction but not maintenance of pairing-induced long-term potentiation. J Neurosci 17:5357–5365.

Palkovits M, Magyar P, Szentágothai J (1971) Quantitative histological analysis of the cerebellar cortex in the cat. III. Structural organization of the molecular layer. Brain Res 34:1–18.

Paninski L, Fellows MR, Hatsopoulos NG, Donoghue JP (2004) Spatiotemporal tuning of motor cortical neurons for hand position and velocity. J Neurophysiol 91:515–532.

Park IM, Meister MLR, Huk AC, Pillow JW (2014) Encoding and decoding in parietal cortex during sensorimotor decision-making. Nat Neurosci 17:1395–1403.

Penfield W, Boldrey E (1937) SOMATIC MOTOR AND SENSORY REPRESENTATION IN THE CEREBRAL CORTEX OF MAN AS STUDIED BY ELECTRICAL STIMULATION. Brain 60:389–443.

Perrett SP, Mauk MD (1995) Extinction of Conditioned Eyelid Responses Requires the Anterior Lobe of Cerebellar Cortex. J Neurosci 15:2074–2080.

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP

(2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature 454:995–999.

Proville RDRD, Spolidoro M, Guyon N, Dugu? GP, Selimi F, Isope P, Popa D, L?na C, Dugué GP, Selimi F, Isope P, Popa D, Léna C (2014) Cerebellum involvement in cortical sensorimotor circuits for the control of voluntary movements. Nat Neurosci 17:1233–1239.

Pugh JR, Raman IM (2006) Potentiation of Mossy Fiber EPSCs in the Cerebellar Nuclei by NMDA Receptor Activation followed by Postinhibitory Rebound Current. Neuron 51:113–123.

Pugh JR, Raman IM (2008) Mechanisms of Potentiation of Mossy Fiber EPSCs in the Cerebellar Nuclei by Coincident Synaptic Excitation and Inhibition. J Neurosci 28:10549–10560.

Raghavan RT, Lisberger SG (2017) Responses of Purkinje cells in the oculomotor vermis of monkeys during smooth pursuit eye movements and saccades: comparison with floccular complex. J Neurophysiol 118:986–1001.

Raymond JL, Lisberger SG, Keck WM (1997) Multiple subclasses of Purkinje cells in the primate floccular complex provide similar signals to guide learning in the vestibulo-ocular reflex. Learn Mem 3:503–518.

Rünger D, Ashby FG, Picard N, Strick PL (2013) Response-mode shifts during sequence learning of macaque monkeys. Psychol Res 77:223–233.

Ryan SB, Detweiler KL, Holland KH, Hord MA, Bracha V (2006) A long-range, wide field-of-view infrared eyeblink detector. J Neurosci Methods 152:74–82.

Safo P, Regehr WG (2008) Timing dependence of the induction of cerebellar LTD.

Neuropharmacology 54:213–218.

Safo PK, Regehr WG (2005) Endocannabinoids control the induction of cerebellar LTD. Neuron 48:647–659.

Schade Powers A, Coburn-Litvak P, Evinger C (2010) Conditioned eyelid movement is not a blink. J Neurophysiol 103:641–647.

Seidler RD, Purushotham A, Kim S-G, Ugurbil K, Willingham D, Ashe J (2005) Neural correlates of encoding and expression in implicit sequence learning. Exp brain Res 165:114–124.

Serapide MF, Zappal? A, Parenti R, Pant? MR, Cicirata F (2002) Laterality of the pontocerebellar projections in the rat. Eur J Neurosci 15:1551–1556.

Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP (2002) Brain-machine interface: Instant neural control of a movement signal. Nature 416:141–142.

Shadmehr R, Smith MA, Krakauer JW (2010) Error Correction, Sensory Prediction, and Adaptation in Motor Control. Annu Rev Neurosci 33:89–108.

Shimansky Y, Saling M, Wunderlich D a, Bracha V, Stelmach GE, Bloedel JR (1997) Impaired capacity of cerebellar patients to perceive and learn two-dimensional shapes based on kinesthetic cues. Learn Mem 4:36–48.

Shin JC, Aparicio P, Ivry RB (2005) Multidimensional sequence learning in patients with focal basal ganglia lesions. Brain Cogn 58:75–83.

Shin JC, Ivry RB (2003) Spatial and Temporal Sequence Learning in Patients with Parkinson's Disease or Cerebellar Lesions. J Cogn Neurosci 15:1232–1243.

Shin S-L, Zhao GQ, Raymond JL (2014) Signals and Learning Rules Guiding Oculomotor Plasticity. J Neurosci 34:10635–10644.

Shinoda Y, Sugihara I, Wu HS, Sugiuchi Y (2000) The entire trajectory of single climbing and mossy fibers in the cerebellar nuclei and cortex. Prog Brain Res 124:173–186.

Siegel JJ, Mauk MD (2013) Persistent Activity in Prefrontal Cortex during Trace Eyelid Conditioning: Dissociating Responses That Reflect Cerebellar Output from Those That Do Not. J Neurosci 33:15272–15284.

Simpson JI, Wylie DR, De Zeeuw CI (1996) On climbing fiber signals and their consequence(s). Behav Brain Sci 19:384–398.

Spencer RMC, Ivry RB (2009) Sequence learning is preserved in individuals with cerebellar degeneration when the movements are directly cued. J Cogn Neurosci 21:1302–1310.

Steinmetz JE, Lavond DG, Thompson RF (1985) Classical conditioning of the rabbit eyelid response with mossy fiber stimulation as the conditioned stimulus. Bull Psychon Soc 23:245–248.

Steinmetz JE, Lavond DG, Thompson RF (1989) Classical conditioning in rabbits using pontine nucleus stimulation as a conditioned stimulus and inferior olive stimulation as an unconditioned stimulus. Synapse 3:225–233.

Steinmetz JE, Rosen DJ, Chapman PF, Lavond DG, Thompson RF (1986) Classical conditioning of the rabbit eyelid response with a mossy-fiber stimulation CS: I. Pontine nuclei and middle cerebellar peduncle stimulation. Behav Neurosci 100:878–887.

Summers JJ, Anson JG (2009) Current status of the motor program: Revisited. Hum Mov Sci 28:566–577.

Suvrathan A, Payne HL, Raymond JL (2016) Timing Rules for Synaptic Plasticity Matched to Behavioral Function. Neuron 92:959–967.

Svensson P, Jirenhed D-A, Bengtsson F, Hesslow G (2010) Effect of Conditioned Stimulus Parameters on Timing of Conditioned Purkinje Cell Responses. J Neurophysiol 103:1329–1336.

Tan H, Gerrits NM (1992) Laterality in the vestibulo-cerebellar mossy fiber projection to flocculus and caudal vermis in the rabbit: A retrograde fluorescent double-labeling study. Neuroscience 47:909–919.

ten Brinke MM, Boele HJ, Spanke JK, Potters JW, Kornysheva K, Wulff P, IJpelaar ACHG, Koekkoek SKE, De Zeeuw CI (2015) Evolving Models of Pavlovian Conditioning: Cerebellar Cortical Dynamics in Awake Behaving Mice. Cell Rep 13:1977–1988.

Thanellou A, Green JT (2011) Spontaneous recovery but not reinstatement of the extinguished conditioned eyeblink response in the rat. Behav Neurosci 125:613–625.

Therrien AS, Bastian AJ (2015) Cerebellar damage impairs internal predictions for sensory and motor function. Curr Opin Neurobiol 33:127–133.

Tkacik G, Prentice JS, Balasubramanian V, Schneidman E (2010) Optimal population coding by noisy spiking neurons. Proc Natl Acad Sci 107:14419–14424.

Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN (2004) A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural

Ensemble, and Extrinsic Covariate Effects. J Neurophysiol 93:1074–1089.

Voicu H, Mauk MD (2006) Parametric analysis of cerebellar LTD in eyelid conditioning. Neurocomputing 69:1187–1190.

Weidemann G, Kehoe EJ (2004) Recovery of the rabbit's conditioned nictitating membrane response without direct reinforcement after extinction. Learn Behav 32:409–426.

Wong K-F, Wang X-J (2006) A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. J Neurosci 26.

Yeo CH, Hardiman MJ, Glickstein M (1985a) Classical conditioning of the nictitating membrane response of the rabbit. I. Lesions of the cerebellar nuclei. Exp brain Res 60:87–98.

Yeo CH, Hardiman MJ, Glickstein M (1985b) Classical conditioning of the nictitating membrane response of the rabbit. II. Lesions of the cerebellar cortex. Exp brain Res 60:99–113.

Yttri EA, Dudman JT (2016) Opponent and bidirectional control of movement velocity in the basal ganglia. Nature 533:1–16.

Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F (2016) Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. Neuron 89:369–383.